

Lesen in beruflichen Handlungskontexten

Anforderungen, Prozesse und Diagnostik



TECHNISCHE
UNIVERSITÄT
DARMSTADT

vom Fachbereich Humanwissenschaften der Technischen Universität Darmstadt

zur Erlangung des Grades Doctor philosophiae (Dr. phil.)

genehmigte Dissertation
von Aileen Deborah Balkenhol
aus Brilon

1. Prüferin: Prof. Dr. Birgit Ziegler (Technische Universität Darmstadt)
2. Prüferin: Prof. Dr. Susan Seeber (Universität Göttingen)

Tag der Einreichung: 28.10.2014
Tag der mündlichen Prüfung: 08.05.2015

Darmstadt 2016

D17

Abstrakt

Lesen ist eine der zentralen Kulturfähigkeiten, die ein Mensch braucht, um in der sogenannten Informationsgesellschaft erfolgreich zu sein. Im Alltag und im Beruf ist Lesen unabdingbar, insbesondere in einer technologisch fortschrittlichen Nation wie Deutschland. Dennoch wird von Betrieben regelmäßig konstatiert und auch in Untersuchungen ermittelt, dass die Lesefähigkeiten vieler Schülerinnen und Schülern kaum den Anforderungen der Berufswelt gerecht werden.

Kernziel dieser Arbeit ist die Entwicklung und Analyse eines diagnostischen Verfahrens zur Erfassung der Lesekompetenz von Auszubildenden. Dabei werden die spezifischen Anforderungsstrukturen beim Lesen in beruflichen Handlungskontexten in den Blick genommen. Dies erfolgt unter besonderer Berücksichtigung der Leseanlässe und der Repräsentationsformate.

Analysen des aktuellen Forschungsstandes zu Leseanforderungen und zur Relevanz von Lesekompetenz im Arbeitskontext ergeben, dass Leseverstehen eine unverzichtbare Kompetenz ist, um in der Mehrheit der Berufe erfolgreich zu sein. Es werden Recherchen im deutschsprachigen wie im anglo-amerikanischen Sprachraum ausführlich erörtert.

Lesen im beruflichen Handlungskontext wird kontrastiert zum Lesen im Lernkontext. Hinsichtlich der vorherrschenden Textgattungen und Repräsentationsformate sind gewisse Unterschiede erkennbar. So dominieren in beruflichen Handlungskontexten Textformen bspw. Instruktionstexte mit formalen Handlungsanweisungen oder konzeptionell mündlich verfasste Anweisungen und Hinweise von Kollegen oder Vorgesetzten. Ferner sind Rechtstexte oder auch non-lineare oder depiktionale Textformen gebräuchlich wie Hypertexte, Tabellen, Formulare, Listen etc. Gegenstand der Untersuchung. Insbesondere auch Hybridformen werden hinsichtlich ihrer Verbreitung und ihrer spezifischen Verarbeitungsmodi betrachtet.

Weiterhin kann angenommen werden, dass im Beruf vorrangig im Handlungskontext einer beruflichen Situation, d.h. zur Bewältigung einer konkreten Aufgabe, gelesen wird. In der vorliegenden Arbeit wird diese These einerseits durch eine curriculare Analyse und durch die Analyse typischer beruflicher Lesesituationen bekräftigt. Es wird ferner herausgearbeitet, auf welcher theoretischen Basis die Verarbeitungsprozesse und insbesondere die Integration multipler multimodaler Repräsentation modelliert werden können.

Die darauf bezogene Entwicklung eines Testinstruments für Auszubildende unterschiedlicher Berufsfelder (gewerblich-technisch, medizinisch pflegend und kaufmännisch-verwaltend) wird skizziert und global analysiert. Zuvor werden alternative Testinstrumente und -verfahren im Hinblick auf ihre Passung bezüglich der herausgearbeiteten theoretischen Anforderungen untersucht. Die Konstruktion eines Testinstruments wird erläutert und anschließend werden die Ergebnisse einer größeren quantitativ-empirischen Studie an berufsbildenden Schulen in Hessen, Niedersachsen und Thüringen bezüglich des Testverfahrens und der Items dargestellt. Sodann wird das Instrument einer ersten Evaluation unterzogen und die Ergebnisse berichtet.

Abstract

Reading is one of the central cultural skills that a person needs to be successful in the so-called information society. In everyday life and at work reading is essential, especially in a technologically advanced country like Germany. Nevertheless, companies and educational researchers notice that the reading skills of pupils have been identified in studies to be inadequate, even though these are regarded as crucial and indispensable basic feature of apprenticeship entry maturity.

This study primarily pursues the aim to investigate the professional read requests and their differing requirement structures for reading-to-learn, i.e., read with the primary aim of acquiring knowledge. Particularly representation formats of the presented information are considered to support the derived from the analysis hypothesis that the cognitive processing processes run on a different basis.

The Analyses present the current state of research to read requests and show the relevance of literacy. Reading comprehension is an essential competence to be successful in the majority of occupations. As part of this analysis, searches are discussed in detail in the German as in the Anglo-American language area.

Here reading to do /reading at work will be examined with a view to dealing with other literary genres and mainly other representation formats as reading to learn. These include at the sides of descriptive textforms e.g. instruction texts with formal procedural instructions or conceptually oral but medial written instructions and guidance from colleagues or superiors. Further legal texts or even non-linear/depiktionale textforms such as hypertext, tables, forms, lists, etc. are subject of the investigation. In particular, hybrid forms are considered in terms of their distribution and their specific processing modes.

Due to the work results, reading at work can be assumed as priority in the context of activity of a professional situation, i.e. to cope with a specific task. This hypothesis is supported by further investigations, a curriculum analysis in the present study and by the analysis of typical occupational reading situations.

In this case, it is worked out on what theoretical basis the manufacturing processes and in particular, the integration of multiple multimodal representation can be modeled.

Furthermore, the development of a measuring instrument for students of vocational schools is outlined and analyzed globally. For this first alternative test are instruments and test procedures examined in terms of their fit relating to the previously worked out theoretical requirements. The design of the test instrument is explained and then the results of a larger quantitative empirical study at vocational schools in Hessen, Lower Saxony and Thuringia with regard to the test method and the Items are shown. Then the instrument is subjected to a first evaluation.

Aileen Balkenhol

Wissenschaftlicher Werdegang

Studium

- 2007-2011 Studium Lehramt an Gymnasien und Gesamtschulen in den Fächern Germanistik und katholische Theologie, Universität Paderborn; Abschluss: 1. Staatsexamen
- 2010-2011 Studium Lehramt am Berufskolleg in den Fächern Germanistik und katholische Theologie, Universität Paderborn; Abschluss: 1. Staatsexamen

Berufstätigkeit

- 11/2014-04/2016 Studienreferendarin an Berufskollegs, ZfsL Hagen, Berufskolleg Olsberg; Fächer: Deutsch und kath. Religionslehre
- 05/2012-10/2014 Wissenschaftliche Mitarbeiterin, Institut für Allgemeine Pädagogik und Berufspädagogik, Arbeitsbereich Berufspädagogik, Technische Universität Darmstadt
- 12/2011-04/2012 Wissenschaftliche Mitarbeiterin, Institut für Erziehungswissenschaft, Arbeitsbereich Berufspädagogik, Universität Paderborn
- 10/2007-11/2011 Studienbegleitendes Praktikum mit den Schwerpunkten Bürokommunikation und Public Relations bei Firma Gussek Haus in Nordhorn und Kamen

Projektteilhabe

- 12/2011-10/2015 „MaK-adapt“ – Studie zur Diagnostik funktionaler Lesekompetenz, BMBF-Drittmittelprojekt

Publikationen

- 2014 Lesekompetenz in der beruflichen Ausbildung und im Berufsalltag. Berufsbildung. Zeitschrift für Praxis und Theorie in Betrieb und Schule, 146, April 2014 (zusammen mit Birgit Ziegler)
- 2012 Diagnostik „funktionaler Lesekompetenz“. *bwp@ Berufs- und Wirtschaftspädagogik – online*, 22, 1-19. [Online](#). (zusammen mit Birgit Ziegler, Christina Keimes, Volker Rexing)

Vorträge

- 03/2014 „Vorgehen und Ergebnisse bei der differentiellen Diagnostik von Lesekompetenz im Lern- und Handlungskontext“
2. Tagung der Gesellschaft für empirische Bildungsforschung in Frankfurt
- 09/2013 „Lesen im beruflichen Handlungskontext“ – Modellierung und Messung „funktionaler“ Lesekompetenz
Herbsttagung der Sektion für Berufs- und Wirtschaftspädagogik der DGfE in Chemnitz
- 03/2013 „Unterscheidung der Diagnostik von Lesekompetenz im Lern- und Handlungskontext – Entwicklung eines Instruments zur Messung von Lesekompetenz im Handlungskontext“
1. Tagung der Gesellschaft für empirische Bildungsforschung in Kiel
- 03/2013 „Lesen im beruflichen Kontext – Diagnostische Anforderungen“ 17. Hochschultage Berufliche Bildung in Essen

Inhalt

Inhalt	1
Abbildungsverzeichnis	2
Tabellenverzeichnis	4
1.....Einleitung und Problemanalyse	7
Beschäftigungshorizont und Aufbau der Arbeit	7
2.....Leseanforderungen in Gesellschaft und Beruf – Probleme und Forschungsstand	12
2.1.1. Relevanz der Lesefähigkeit für Individuum und Gesellschaft	12
2.1.2. Leseanforderungen und -anlässe in der beruflichen Ausbildung und im Berufsalltag	18
2.1.3. Vergleich von Anspruch und Leistung beruflicher Lesekompetenz	35
2.1.4. Leseförderung in der beruflichen Bildung	38
2.1.5. Rekapitulation des Untersuchten I	39
3.....Theoretische Ausarbeitungen	40
3.1. Lesen – Begriff, Situation, Material, Modellierung und Messung	40
3.1.1. Annäherung an das Konstrukt „Lesekompetenz“	40
3.1.2. Kommunikationssituation und Leseintention	49
3.1.3. Analyse des spezifisch beruflichen Lesematerials	57
3.1.4. Leseprozess-theorien/Lesemodelle	72
3.1.5. Aspekte der Textverständlichkeitsforschung und ihre Übertragbarkeit auf depiktionale Repräsentationen	115
3.1.6. Lesekompetenzmessung – Stand, Testverfahren und Passung auf den beruflichen Kontext	123
3.1.7. Rekapitulation des Untersuchten II	134
4.....Ansätze zur Entwicklung eines Instruments zur Untersuchung des Einflusses der Repräsentationsformate und kognitiven Anforderungsbereiche	136
4.1. Entwicklung des Testinstruments	136
4.1.1. Überlegungen zum Testformat	136
4.1.2. Rahmenbedingungen der Test- und Itementwicklung	138
4.1.3. Theoretisches Kompetenzmodell	142
4.1.4. Umsetzung	148
4.2. Analyse des Testinstruments	161
4.2.1. Kalibrierung der Testitems – empirische Schwierigkeitsbestimmung an Hand der Lösungswahrscheinlichkeit	161
4.2.2. Exkurs I – Einflüsse des Präsentationsmediums	181

4.2.3. Exkurs II – Analyse der Itemschwierigkeit unter Aspekten der Verständlichkeitsforschung I	196
4.2.4. Exkurs III – Analyse der Itemschwierigkeit unter Aspekten der Verständlichkeitsforschung II	212
4.2.5. Exkurs IV – Vergleich der Verfahren zur Schwierigkeitsbestimmung	215
4.2.6. Niveaustufenmodellierung	216
5.....Praxistransfer, Rekapitulation & Ausblick	222
5.1. Praktischer Einsatzbereich des Tests und Evaluation	222
5.2. Beantwortung der Forschungsfragen & Reflexion des Untersuchten	227
5.3. Ausblick und Anschlussforschung	229
6.....Literatur	232
7.....Anhang	254
7.1. Ratingbogen	255
7.2. Ergänzende Tabellen und Abbildungen	256
8.....Eidesstattliche Erklärung	269

Abbildungsverzeichnis

Abbildung 1 – Determinanten des Leseprozesses (Jenkins 1979)	48
Abbildung 2 – Determinanten des Leseprozesses II.....	49
Abbildung 3 – Interpretationskomponenten der Kommunikation mit Instruktionstexten (Saile 1982. S. 135).....	64
Abbildung 4 – Weiterentwicklung der Interpretationskomponenten der Kommunikation mit Instruktionstexten.....	65
Abbildung 5 – Kommunikationszusammenhang Autor-Leser (Ziegler 2013 in Anlehnung an Schnotz 1994)	73
Abbildung 6 – Modell der DCT	110
Abbildung 7 – Schematische Darstellung eines integrierten Modells des Text- und Bildverstehens (vgl. Schnotz/Bannert 1999; Schnotz/Dutke 2004)	112
Abbildung 8 – Kompetenzmodell des LTB3.....	128
Abbildung 9 – adaptiver Testverlauf (Frey (2012))	141
Abbildung 10 – Lesen in beruflichen Handlungskontexten (Modellierung angelehnt an Schnotz/Bannert bzw. Schnotz/Dutke)	144

Abbildung 11 – Repräsentationsformate und Anforderungsklassen	146
Abbildung 12 – Item 60 Seite 1	150
Abbildung 13 – Item 60 Seite 2	151
Abbildung 14 – Item 59	152
Abbildung 15 – Item 6 Seite 1	153
Abbildung 16 – Item 6 Seite 2	153
Abbildung 17 – Item 49 Seite 1	154
Abbildung 18 – Item 49 Seite 2	155
Abbildung 19 – Item 64 Seite 1	156
Abbildung 20 – Item 64 Seite 2	157
Abbildung 21 – Item 62 Seite 1	159
Abbildung 22 – Item 62 Seite 2	160
Abbildung 23 – Domänenfolge der Testhefte (Bernhard 2012)	165
Abbildung 24 – Latent Distribution Map (CBA)	167
Abbildung 25 – Test Characteristic Curve (CBA)	168
Abbildung 26 – Test Information Function (CBA).....	168
Abbildung 27 – ICC Item 6 (CBA).....	173
Abbildung 28 – ICC Item 59 (CBA).....	173
Abbildung 29 – ICC Item 60 (CBA).....	174
Abbildung 30 – ICC Item 49 (CBA).....	174
Abbildung 31 – ICC Item 64 (CBA).....	175
Abbildung 32 – ICC Item 62 (CBA).....	175
Abbildung 33 – TCCs (CBA + PP)	187
Abbildung 34 – TIF (CBA + PP)	187
Abbildung 35 – Item 1 Seite 1	190
Abbildung 36 – Item 1 Seite 2	191
Abbildung 37 – Item 17 Seite 1	192

Abbildung 38 – Item 17 Seite 2	192
Abbildung 39 – ICCs Item 6 (CBA + PP)	194
Abbildung 40 – IIFs Item 6 (CBA + PP)	194
Abbildung 41 – ICCs Item 62 (CBA + PP)	195
Abbildung 42 – ICCs Item 64 (CBA + PP)	195
Abbildung 43 – Reliabilitätsentwicklung in Abhängigkeit der Itemanzahl (Simulation) (Frey 2013)	223
Abbildung 44 - Motivation der Probanden (Selbstauskunft)	224
Abbildung 45 – Antwortverhalten Relevanz der Inhalte	226
Abbildung 46 – IIF Item 6 (CBA)	260
Abbildung 47 – IIF Item 49 (CBA)	260
Abbildung 48 – IIF Item 59 (CBA)	261
Abbildung 49 – IIF Item 60 (CBA)	261
Abbildung 50 – IIF Item 62 (CBA)	261
Abbildung 51 – IIF Item 64 (CBA)	262
Abbildung 52 – ICC Item 6 (PP)	263
Abbildung 53 – ICC Item 62 (PP)	263
Abbildung 54 – ICC Item 64 (PP)	264
Abbildung 55 – IIF 6 (PP)	264
Abbildung 56 – IIF Item 62 (PP)	265
Abbildung 57 – IIF Item 64 (PP)	265
Abbildung 58 – IIFs Item 62 (CBA + PP)	266
Abbildung 59 – IIFs Item 64 (CBA + PP)	266

Tabellenverzeichnis

Tabelle 1 – exemplarische Leseanlässe der beruflichen Ausbildung gemäß Lehrplan (eigene Darstellung)	28
Tabelle 2 – Taxonomie der Leseintentionen	55

Tabelle 3 – Kategorisierung von Lesematerial	97
Tabelle 4 – Orthogonal Relationship between Mental Codes and Sense Modalities (Sadoski/Paivio 2004, S. 5)	108
Tabelle 5 – Itempool (gesamt) Matrixdarstellung	149
Tabelle 6 – Itemkennwerte (CBA)	169
Tabelle 7 – Itemkennwerte Item 6, 49, 59, 60, 62 und 64	172
Tabelle 8 – Antwortauswahl Item 6 (CBA)	176
Tabelle 9 – Einordnung der Antworten für Item 6 (CBA)	176
Tabelle 10 – Antwortauswahl Item 49 (CBA)	177
Tabelle 11 – Einordnung der Antworten für Item 49 (CBA)	177
Tabelle 12 – Antworten Item 60 (CBA)	178
Tabelle 13 – Einordnung der Antworten für Item 60 (CBA)	178
Tabelle 14 – Antworten Item 64 (CBA)	178
Tabelle 15 – Einordnung der Antworten für Item 64 (CBA)	179
Tabelle 16 – CBA Werte aufgeschlüsselt nach Repräsentationsformaten	180
Tabelle 17 – CBA Werte aufgeschlüsselt nach kognitiven Anforderungsbereichen	180
Tabelle 18 - Stichproben (Kalibrierung, Pilotierung CAT, Pilotierung PP)	182
Tabelle 19 – Itemkennwerte PP	183
Tabelle 20 – PP Werte aufgeschlüsselt nach Repräsentationsformaten	184
Tabelle 21 – PP Werte aufgeschlüsselt nach kognitiven Anforderungsbereichen	184
Tabelle 22 – Itemkennwerte Items 6, 62 und 64 (PP)	185
Tabelle 23 – Antwortauswahl Item 64 (PP)	186
Tabelle 24 – Einordnung der Antworten für Item 64 (PP)	186
Tabelle 25 – Mittelwert der Estimates (PP und CBA–PP-Items) nach Repräsentationsformaten	189
Tabelle 26 – Mittelwert der Estimates (PP und CBA–PP-Items) nach Anforderungsbereichen	189
Tabelle 27 – Vergleich der Itemkennwerte CBA und PP für die Items 6, 62 und 64	193

Tabelle 28 – Raterreliabilitäten der schwierigkeitsbestimmenden Merkmale	201
Tabelle 29 – Bewertung der Kappa-Statistiken nach Landis/Koch (1977)	203
Tabelle 30 – Intraraterreliabilität Rater A	204
Tabelle 31 – Intraraterreliabilität Rater B	204
Tabelle 32 – Abweichungen der Einschätzungskategorien Rater B	205
Tabelle 33 – Punkteverteilung der Items nach dem Rating	205
Tabelle 34 – Bewertung des Korrelationskoeffizienten nach Brosius (1998)	207
Tabelle 35 – Korrelation der Punktwerte der Rater	207
Tabelle 36 – Raterreliabilitäten Itemebene	208
Tabelle 37 – Vergleich Extremwerte Rater A und B	211
Tabelle 38 – Ratingergebnisse und Raterreliabilitäten nach Repräsentationsformaten	211
Tabelle 39 – Ratingergebnisse und Raterreliabilitäten nach kognitiven Anforderungsbereichen	212
Tabelle 40 – Mittelwerte der Estimates der approximierten schwierigkeitsbestimmenden Merkmale	213
Tabelle 41 – Korrelationen der Schwierigkeiten Rating + CBA Kalibrierung	216
Tabelle 42 – Leseanlässe im Rahmenlehrplan für die berufliche Grundbildung des Ausbildungsberufs Maurer/in (Auszug) (Keimes/Rexing 2011b, S. 4)	256
Tabelle 43 – Übersicht über relevantes Textmaterial (Experteninterviews) (Keimes 2014, S. 154)	257
Tabelle 44 – Textrezeptive Handlungsfelder (Keimes 2014, S. 164)	258
Tabelle 45 – Einordnung der Texte in das Modell funktionale Lesekompetenz	259
Tabelle 46 – Itemkennwerte (CBA [blau] + PP)	267
Tabelle 47 – Ergebnisse CBA Kalibrierung für die im PP gelaufenen Items aufgeschlüsselt nach Anforderungsbereichen und Repräsentationsformaten	268

1. Einleitung und Problemanalyse

Beschäftigungshorizont und Aufbau der Arbeit

Der englische Arzt Thomas Sydenham (1624-1689) machte im 17. Jahrhundert eine ebenso simple wie zu dem Zeitpunkt revolutionäre Feststellung, die im Wesentlichen bis zum heutigen Tag gültig geblieben ist: *Am Anfang steht die Diagnose und jede weitere Handlung ist weitgehend von ebendieser Diagnose bestimmt.* Auch wenn Sydenham diese Aussage in Bezug auf Krankheiten traf, kann sicherlich davon ausgegangen werden, dass ein Prozess zuerst zu beschreiben, zu erkennen und verstehen ist, bevor in ihn zielgerichtet und erfolgreich eingegriffen werden kann – gleiches sollte sodann für den komplexen kognitiven Prozess des Lesens gelten. Zunächst ist zu verstehen, wie der Prozess abläuft, wie er funktioniert, welche Einflüsse auf ihn wirken und welche Schwierigkeiten auftreten können, um dann bei Bedarf angemessene Optimierungsstrategien anbieten zu können. Es gibt verschiedene Gründe, warum eine Beschäftigung mit dieser Thematik und daher auch mit dieser Arbeit notwendig oder gar lohnend ist:

Lesen ist eine der zentralen Kulturfähigkeiten, die ein Mensch braucht, um in der sogenannten Informationsgesellschaft erfolgreich zu sein. Im Alltag und im Beruf ist Lesen unabdingbar, insbesondere in einer technologisch fortschrittlichen Nation wie Deutschland (vgl. Kapitel 2).

Dennoch stellen Betriebe und Bildungsforscher¹ fest, dass in bestimmten Sparten des beruflichen Bildungssystems die Bewerber nicht über ausreichende Fähigkeiten im Bereich der Sprachbeherrschung verfügen (vgl. Gartz et al. 1999, S. 132). Daneben wurde die Lesefähigkeiten in Studien als unzureichend ermittelt (vgl. Lehmann et al. 2005), obwohl diese als entscheidendes und unverzichtbares Basismerkmal der Ausbildungsreife gelten (vgl. Bundesagentur für Arbeit 2009, S. 24f). So gibt beispielsweise in einer Umfrage ein knappes Drittel der befragten Betriebe als Grund für ihre fehlende Ausbildungsbereitschaft die Bewerbersituation – genauer den Mangel an kompetenten Bewerbern – an (vgl. Gartz et al. 1999, S. 18). Ferner kann konstatiert werden, dass ca. 15% der Lehrstellenbewerber abgelehnt werden, da ihre Lese- und Schreibfähigkeiten unzureichend sind (vgl. Becker-Mrotzek et al. 2006, S. 6). Betrachtet man sodann das Verständnis der Betriebe von notwendiger Kompetenz in Bezug auf die Fähigkeiten der Bewerber, so stellt man fest, dass seitens der Betriebe keine ungewöhnlichen Forderungen gestellt werden. Stattdessen werden immer wieder, bspw. im Abstand von 10 Jahren, die sogenannten Basisqualifikationen, wie beispielshalber basale Lesefähigkeiten betont (vgl. Gartz et al. 1999, S. 18–21; Bundesagentur für Arbeit 2009, S. 22–30).

Die in der Arbeit folgende Analyse und Zusammenfassung des aktuellen Forschungsstandes zu Leseanforderungen – unter besonderer Berücksichtigung des beruflichen Kontextes – und zur beruflichen wie gesellschaftlichen Relevanz von Lesekompetenz lässt den Schluss zu, dass Leseverstehen eine unverzichtbare Kompetenz zu sein scheint, um in der Mehrheit der Berufe erfolgreich zu sein. Weiterhin sind die Leseanlässe im Beruf in den letzten Jahren nicht weniger

¹ „Verbum hoc ‚si quis tam masculos quam feminas com plectitur‘ (Corpus Iuris Civilis Dig. L, 16,1)

geworden (vgl. Kapitel 2.1). Um zu dieser Analyse zu kommen, wurden Recherchen im deutschsprachigen wie im anglo-amerikanischen Sprachraum angestellt, welche zunächst das Ergebnis erbrachten, dass Lesen oder Leseanforderungen in beruflichen Kontexten wenig erforscht zu sein scheint. Insbesondere neuere und deutschsprachige Ergebnisse fehlen, so dass größtenteils auf amerikanische Forschung der 80er Jahre zurückgegriffen werden musste.

Lesen im Beruf scheint dabei einerseits den Umgang mit anderen Textgattungen und auch überwiegend anderen Repräsentationsformaten als das Lesen im Kontext des Wissenserwerbs zu erfordern (vgl. Ziegler et al. 2012). Dazu gehören auf Seiten der deskriptiven Textformen bspw. Instruktionstexte mit formalen Handlungsanweisungen (u.a. Betriebsanleitungen, Herstellerhinweise) oder konzeptionell mündlich verfasste Anweisungen und Hinweise von Kollegen oder Vorgesetzten. Ferner sind Rechtstexte wie z.B. Vorschriften, Verträge und AGBs zu nennen oder auch non-lineare Textformen wie Hypertexte, Tabellen, Formulare, Listen etc. Es finden aber auch verstärkt Formen depiktionaler Natur Anwendung wie bspw. Grafiken, Schaubilder, Abbildungen, und Diagramme (vgl. Schnotz 2005). Häufig treten diese Formatkategorien nicht in Reinformen, sondern in Mischformen auf. So sind z.B. in Betriebsanleitungen auch oft erklärende/ergänzende Abbildungen oder in Berichten auch veranschaulichende Grafiken zu finden (vgl. Küster 1982; Andersen 2009; Keimes 2014).

Dabei gilt es zusätzlich zu berücksichtigen, dass im Zeitalter der multimedialen Informationspräsentation und der Omnipräsenz von elektronischer Informationstechnik Lesen nicht mehr nur linearen, deskriptiven, gedruckten Texten Informationen zu entnehmen heißt. Stattdessen bedeutet es, sich in non-linearen Hypertextstrukturen unterschiedlichster medialer Repräsentationen zurechtzufinden und deskriptionale sowie depiktionale Informationen zu dechiffrieren, zu verknüpfen und weiter zu verwenden, ggf. auch aus zwei getrennten Informationsteilen eine neue Information zu generieren.

Andererseits darf auf Grund der vorliegenden Befunde (vgl. Kap. 2.2) angenommen werden, dass im Beruf vorrangig im Handlungskontext einer beruflichen Situation, d.h. zur Bewältigung einer konkreten Aufgabe, gelesen wird. Diese These wird einerseits durch erste weitergehende Untersuchungen (vgl. Keimes 2014), eine curriculare Analyse in der vorliegenden Arbeit und durch die Analyse typischer beruflicher Lesesituationen (vgl. Mikulecky 1988) gestützt. Eine Studie von Schmalhofer/Glavanov (1986) gibt ferner Anlass zur Annahme, dass Texte anders rezipiert werden, wenn es darum geht, den zu erschließenden Inhalt in eine Handlung umzusetzen (vgl. Schmalhofer/Glavanov 1986). Daher darf angenommen werden, dass der Prozess des Transfers einer schriftlich fixierten Information in eine Handlung es erfordert aus dem Gelesenen ein mentales Modell zu bilden bzw. das Gelesene in ein bereits durch andere Texte, gemachte Erfahrungen, mündliche Erklärungen oder durch bereits vorhandenes Weltwissen generiertes mentales Modell der erforderlichen Handlung zu integrieren (vgl. Kintsch 1974; Kieras/Bovair 1984).

Die vorliegende Arbeit verfolgt daher primär folgendes Ziel:

Es sollen die allgemeinen beruflichen Leseanforderungen und ihre differierenden Anforderungsstrukturen zum Lesen im Lernkontext, d.h. Lesen mit dem Primärziel des Wissenserwerbs, unter besonderer Berücksichtigung der Repräsentationsformate der dargebotenen Informationen untersucht werden, um die aus der Analyse abgeleitete These zu stützen, dass die kognitiven Verarbeitungsprozesse auf einer anderen Basis ablaufen.

Dabei soll herausgearbeitet werden, auf welcher theoretischen Basis die Verarbeitungsprozesse und insbesondere die Integration multipler multimodaler Repräsentation modelliert werden können.

Nachgeordnet soll die Entwicklung eines Testinstruments für Schüler der berufsbildenden Schulen – welches die oben beschriebenen Problembereiche, die bisher in derartigen Lesetests nicht im Fokus standen, bewusst aufgreift – skizziert und global analysiert werden. Dazu werden zunächst alternative Testinstrumente und -verfahren im Hinblick auf ihre Passung bezüglich der zuvor herausgearbeiteten theoretischen Anforderungen untersucht, die Konstruktion eines Testinstruments erläutert und anschließend die Ergebnisse einer größeren quantitativ-empirischen Studie an berufsbildenden Schulen in Hessen, Niedersachsen und Thüringen bezüglich des Testverfahrens und der Items dargestellt werden.

Die Arbeit ist dazu wie folgt strukturiert:

Sie besteht aus drei Abschnitten, wovon der erste eine allgemeine Einführung und die Problemanalyse enthält. Hierbei werden in Kapitel 2 zunächst die sprachlichen, textsystematischen und situativen Anforderungen, die Schülerinnen und Schüler innerhalb der beruflichen Ausbildung und im späteren Arbeitsleben zu erwarten haben, untersucht.

Zu Beginn werden in Kap. 2.1 die gesellschaftliche und ökonomische Bedeutung der Lesefähigkeit aufgearbeitet und insbesondere auf den beruflichen Kontext bezogen geklärt, welche möglichen Konsequenzen unzureichende Lesekompetenz haben kann.

In Kapitel 2.1.2 werden weiterhin zunächst die Leseanforderungen und -anlässe in der beruflichen Ausbildung und im Berufsalltag kurz an Hand von Studien zusammengefasst und herausgearbeitet, welche Textsorten und welche Menge an Leseaufgaben zu bewältigen sind. Dabei zeigt sich, dass das Lesen im Beruf sowohl depiktionale als auch deskriptionale und vor allem Mischformen als Präsentationsformate kennt. Daneben wird offenkundig, dass es kaum Berufe gibt, die nur sehr geringe Leseanteile haben und nahezu keine Berufe gibt, in denen gar nicht gelesen werden muss. Dabei werden auch gesondert die Leseanforderungen im beruflichen Kontext betrachtet, da diese sich entscheidend von denen des Lesens zum Wissenserwerb unterscheiden.

Nach der Problemanalyse werden im 3. Abschnitt die theoretischen Grundlagen und Ausarbeitungen vorgestellt. Es wird versucht zu beantworten, was die spezielle kognitive Tätigkeit des Lesens auszeichnet und wieso dafür ggf. eine gesonderte Facette der Kompetenz benötigt wird.

Daran anknüpfend werden schließlich verschiedene Erklärungsmodelle bezüglich des Leseprozesses und der Verstehenstiefe kurz vorgestellt, bevor im Kapitel 3.1.6 zentrale Ansätze zur Lesekompetenzmessung dargestellt werden.

Für den Bereich der Lesekompetenz gibt es nicht erst seit dem „PISA-Schock“ – dem auffallend schlechten Abschneiden deutscher Schülerinnen und Schüler bei der ersten PISA-Studie 2001 – zahlreiche Testverfahren (vgl. Artelt et al. 2007, S. 6). Jedoch sind diese primär für den Einsatz im Primarbereich und in der Sekundarstufe I konzipiert. Dabei wird auch deutlich, dass diese Ansätze zwar die unterschiedlichen Repräsentationsmodi berücksichtigen, diese jedoch meist nicht als strukturelle Dimension, die Auswirkungen auf die ablaufenden kognitiven Prozesse haben können, betrachten und diese gezielt untersuchen. Es wurde zwar bereits versucht inhaltliche Dimensionen von Beruflichkeit einzubeziehen, dessen ungeachtet wurde jedoch die genuin andersgearteten Leseanforderungen – eine Handlung auszuführen oder fehlende Informationen in ein bestehendes Schema zu ergänzen – außen vor gelassen. So kann gezeigt werden, dass eine grundsätzliche Berücksichtigung dieser beiden Aspekte sowohl bei der Konstruktion als auch der Analyse von (berufsbezogenen) Lesetests bisher noch nicht ausführlich erfolgt ist.

Daran anknüpfend wird in Kapitel 4 die Testentwicklung beschrieben. Dazu wird knapp auf allgemeine Aspekte der Testkonstruktion bzw. des Testformats eingegangen (vgl. Kap. 4.1.1), um dann die speziellen Bedingungen der Entwicklung der Testfragen (vgl. Kap. 4.1.2) neben der Beschreibung des theoretischen Kompetenzmodells (vgl. Kap. 4.1.3) sowie einige exemplarische Konstruktionen darzustellen (vgl. Kap. 4.1.4). Eine der wichtigsten Vorgaben² dabei war, dass die Items im MC Format computerisiert vorzuliegen hatten. Dabei sollten sowohl die Bearbeitungszeit wie auch die Itemlänge nach Möglichkeit kurz ausfallen.

Der nächste große Abschnitt der Arbeit widmet sich daran anschließend der empirischen Auswertung der erhobenen Daten. Die entwickelten Fragen werden sowohl einer qualitativen Analyse unter Aspekten der Textverständlichkeitsforschung unterzogen, die Ergebnisse eines auf diesen theoretischen Aspekten basierenden Ratings diskutiert, als auch die Ergebnisse der quantitativen Erhebungen vorgestellt. Die Ergebnisse beider Schwierigkeitseinschätzungen werden anschließend verglichen und analysiert – berücksichtigt und untersucht wird auch der mögliche Einfluss des Präsentationsmedium (Paper-Pencil vs. computerisiert).

Dazu erfolgt in Kapitel 4.2.1 zunächst eine ausführliche Darlegung des Einsatzes der entwickelten Testfragen im Rahmen einer Kalibrierungs- und einer Pilotierungsstudie für das BMBF Projekt „Messung allgemeiner Kompetenzen – adaptiv (MaK-adapt)“³.

Die Items wurden auf Basis einer Stichprobe ($n > 1500$) bestehend aus Schülerinnen und Schülern kurz vor Ende ihrer Berufsausbildung kalibriert, d.h. die Schwierigkeiten wurden unter

² Die Items wurden im Rahmen eines von der Forschungsinitiative ASCOT (BMBF) geförderten Verbundprojekts (Messung allgemeiner Kompetenzen – adaptiv (MaK-adapt) am Standort Darmstadt entwickelt (Details siehe Kap. 4.1.2).

³ FKZ 01DB1103

Einsatz eines Raschmodells gemäß der Item Response Theory skaliert. Die sich daraus ergebenden Werte sind die Annahme der quantitativen empirischen Schwierigkeit der Items. Während der Kalibrierungsstudie wurden durch Beobachtungen Probleme der Probanden mit den entwickelten Testfragen (Aufgabenstellung, Wortwahl, Antwortmöglichkeiten, Unvertrautheit des präsentierten Materials,...) ermittelt, auch wurden die Items im Anschluss hinsichtlich möglicher Diff-Effekte untersucht und sodann ggf. angepasst oder ausgesondert, so dass u.a. materielle Bearbeitungsprobleme oder Diff-Effekte als schwierigkeitsbeeinflussender Faktor nach Möglichkeit ausgeschlossen werden können.

Weiterhin wurden verschiedene Exkurse unternommen, um die Items von möglichst verschiedenen Seiten zu beleuchten.

In einer zweiten Studie (Pilotierungsstudie mit einem gesamt $n > 1500$) wurden die beibehaltenen Items einem Teil der Stichprobe ($n=500$) als Paper-Pencil-Version vorgelegt, um sich dem Einfluss des Präsentationsmediums auf das Antwortverhalten annähern zu können. So können Hinweise auf medial-präsentationsbezogene Unterschiede bei der Bearbeitung der Items untersucht werden. Die Ergebnisse dieser Studie sind in Kapitel 4.2.2 zu finden.

Darauffolgend erfolgt eine Analyse der Itemschwierigkeit unter Aspekten der Verständlichkeitsforschung. Bei dieser Analyse wird einerseits der entwickelte Kriterienbogen vorgestellt, der die abgeleiteten einflussreichsten schwierigkeitsbestimmenden Parameter aus der Textverständlichkeitsforschung und eigener Ansätze zur analogen Klassifikation ebendieser Parameter für depiktionale Präsentationsformate enthält. An Hand dieser Kriterien wurden nach einer Schulung alle Items von zwei unabhängigen Ratern klassifiziert. Die Ergebnisse der Einschätzung und die Interrater-Reliabilität sowie weitere Analysen der Einschätzungen werden in Kapitel 4.2.3 dargelegt. Die Rating Kategorien und ihre tatsächlicher Einfluss auf die Schwierigkeit der Items werden in Kapitel 4.2.4 einer kleinen empirischen Untersuchung unterzogen, in der die Parameter auf ihre tatsächliche Erklärungsmacht hin analysiert werden. Abschließend zu diesen Betrachtungen werden die durchgeführten Verfahren zur Schwierigkeitsbestimmung der Items kurz miteinander verglichen (vgl. Kap. 4.2.5)

An Hand dieser Daten lassen sich diverse Analysen (Repräsentationsformate, Anforderungskontexte...) durchführen. Deren Ergebnisse werden ebenfalls Gegenstand der Arbeit sein.

Auf dieser Basis werden anschließend in Kapitel 4.2.6 Ansätze zur Niveaustufenmodellierung gezeigt und Kompetenzniveaus modelliert und beschrieben.

Abschließend werden in Kapitel 6 die praktisch gemachten Erfahrungen beim Einsatz der Testfragen kurz reflektiert, um in Kapitel 7 ein Fazit zu ziehen und in Kapitel 8 einen kleinen Ausblick auf weitere Forschungsdesiderata geben zu können.

2. Leseanforderungen in Gesellschaft und Beruf – Probleme und Forschungsstand

2.1.1. Relevanz der Lesefähigkeit für Individuum und Gesellschaft

Lesen zu können ist einer der größten Einflussfaktoren auf die Gestaltungsmöglichkeiten des gesellschaftlichen und beruflichen Lebens in Deutschland im 21. Jahrhundert. Damit ist nicht die rein formale Beherrschung der basalen Stufen der Kulturtechnik gemeint. Diese reicht als einfache Grundqualifikation nicht mehr aus. Sie muss durch die Fähigkeit ergänzt werden, Informationen – die in schriftlicher Form vorliegen – zu verstehen und weiterzuverarbeiten (vgl. Husfeldt 2001, S. 12).

Diese elaborierte Stufe der Lesekompetenz spielt bei der steigenden Vielzahl an Ansprüchen, die an das Bildungsniveau der Bürger gestellt werden, eine fundamentale Rolle. Auch wenn es seit den 60er Jahren des 20. Jahrhunderts Theorien gibt, dass durch die Entwicklung und Verbreitung von Telefon und Television, orale und visuelle Kommunikationsformen das geschriebene Wort zukünftig ersetzen würden (vgl. Birkerts 1994), dass das Lesen als gesellschaftliche relevante Kommunikationsform gänzlich verschwinden würde, dass Bilder, die ohne spezielle Bildungsvoraussetzungen zugänglich seien, die Schriftsprache ersetzen würden (vgl. Hurrelmann 2008, S. 18)⁴, hat das geschriebene Wort seinen Platz als wichtigstes Medium der Kommunikation bislang nicht eingebüßt (vgl. Kirsch et al. 2002a, S. 15). Im Gegenteil: Television konnte sich nicht als dominantes Bildungsmedium etablieren, Computer haben das Lesen gerade nicht überflüssig gemacht, sondern in multimediale Zusammenhänge integriert. Die Kommunikationsformen haben sich dabei also nicht substituiert, sondern seit den 80er Jahren weiter ausdifferenziert, wurden vielfältiger und komplexer (vgl. Bungarten 1988), wobei Lesen abhängig vom medialen Kontext verschiedener Couleur sein kann (vgl. Hurrelmann 2008, S. 19).

Hurrelmann geht so weit die folgende These aufzustellen: „Lesekompetenz ist zum unverzichtbaren Teil einer umfassenden Medienkompetenz geworden“ (Hurrelmann 2008, S. 19).

Diese Bedeutung spiegelt sich unter anderem darin, dass Lesefertigkeiten einen entscheidenden Einfluss auf den schulischen Erfolg in den beiden Sekundarstufen haben (vgl. Holloway 1999). Nach Groeben/Hurrelmann (2004) ist Lesen sogar die zentrale Bedingung für die Entwicklung weiterer „transmedialer Kompetenzdimensionen“ (Groeben/Hurrelmann 2004, S. 449). Lesekompetenz ermöglicht somit die Nutzung von schriftlichen Informationen als Ressourcen und verschafft Zugang zu Institutionen, die konstituierend schriftlich kommunizieren (Behörden, Ämter und Universitäten) (vgl. Olson 1994). Elwert (2001) hat dazu ein Konstrukt sogenannter „gesellschaftlicher Lesekompetenz“ („societal literacy“) (vgl. Elwert 2001, S. 54) entwickelt, welches die Bedeutung der Lesekompetenz in der modernen bürokratischen Gesellschaft unterstreicht, denn im Rechtswesen, im Geschäftsverkehr und in der Forschung werden Schriftstücke

⁴ Was im Übrigen eine interessante Annahme ist unter der Perspektive, dass dies zur Zeiten geringer Alphabetisierungsraten (bspw. im europäischen Mittelalter) durchaus eine übliche Kommunikationsart darstellte.

und schriftlich festgelegte Prozeduren verwendet (bspw. Gesetze, Verträge und Veröffentlichungen), die verstanden werden müssen, um in diesen Bereichen agieren und reagieren, sowie erfolgreich eigenen Ziele verfolgen zu können.

Die weitergehende Annahme, dass Lesen eine Notwendigkeit sei, um einerseits politisch, gesellschaftlich und beruflich kompetent handeln zu können und, um so einflussreich zu sein wie subjektiv erwünscht oder nach externen Kriterien notwendig und um die eigene berufliche wie gesellschaftliche Position zu verbessern, könnte sich daraus ableiten. Dies nimmt auch Power (1983) an:

„If there is a clear relation between level of literacy and quality of life, there is an equally clear connection between the distribution of literacy and the possession of power. From the earliest times, literate members of society have had disproportionate power: priests, oracles, poets and medicine men have all shared as the source of their power a certain literacy, and their modern counterparts – professors, lawyers, engineers and doctors – continue to exercise power, in part because of their control of language.”⁵ (Power 1983, S. 24)

Krahn und Lowe führen dazu aus:

“The ability to read, write and use numerical information is crucial for labor market success and social well-being. Inadequate literacy skills reduce employment prospects and limit participation in society”⁶ (Krahn/Lowe 1999, S. 38) und belegen dies durch Verweise auf Power (1983), Centre for Educational Research and Innovation (CERI) (1992), OECD (1995) (vgl. Krahn/Lowe 1999, S. 38).

Die PISA-Studie nimmt an, dass Lesen eine der Kompetenzen ist, die in modernen Gesellschaften für eine zufriedenstellenden Lebensführung, was sowohl den persönlichen, den sozialen als auch den ökonomischen Bereich betrifft, notwendig sei (vgl. Artelt/Stanat 2010, S. 353). Auch die Internationale Erhebung über Grad und Verteilung elementarer Grundqualifikationen Erwachsener (IALS) (vgl. OECD/Statistics Canada 1997) und eine Studie von Kirsch, Jungblut, Jenkins und Kolstad (2002b) weisen darauf hin, dass mangelnde Lesekompetenz im Zusammenhang mit kriminellen Verhalten oder dem Bedarf an staatlicher Unterstützung stehen kann. Weiterhin kommen sie zu der Schlussfolgerung, dass mangelnde Lesefähigkeit in der westlichen Gesellschaft eine geringere Chance auf eine Vollzeitbeschäftigung bedeute. Das rühre daher, dass Lesen die Basis für den Erwerb von Wissen und damit für das erfolgreiche Abschneiden in

⁵ „Wenn es einen klaren Zusammenhang zwischen dem Grad der Lesekompetenz und der Lebensqualität gibt, dann existiert eine ähnlich klare Verbindung zwischen der Verteilung der Lesekompetenz und dem Besitz von Macht. Von den frühesten Zeiten an hatten gebildete Mitglieder der Gesellschaft unverhältnismäßig viel Macht: Priester, Propheten, Dichter und Medizinmänner teilten alle als Quelle ihrer Macht eine gewisse Lesekompetenz und ihre modernen Gegenstücke – Professoren, Rechtsanwälte, Ingenieure und Ärzte – fahren weiterhin fort, zum Teil auf Grund ihrer Kontrolle der Sprache, Macht auszuüben.“

⁶ „die Fähigkeit zu lesen, zu schreiben und numerischen Informationen zu benutzen ist entscheidend für den Erfolg auf dem Stellenmarkt und für gesellschaftliches Wohlbefinden. Unangemessene Lesefähigkeiten reduzieren Einstellungsaussichten und begrenzen die Teilnahme an der Gesellschaft“

allen Schulfächern sei, was bspw. die PISA Studie (vgl. Baumert et al. 2001a) auch für Deutschland belegen kann.

Der Zusammenhang von Lesen und Beruf⁷ ergibt sich auch aus der Feststellung, die Pätzold trifft: „Immer mehr Berufe sind sogenannte Kommunikationsberufe, d.h. die Kommunikationsfähigkeit macht den größten Teil der beruflichen Tätigkeit aus, gleichgültig um welche hierarchischen Ebenen es sich handelt“ (Pätzold 2010, S. 162). Daraus schließt er weiter, dass die hohen Abbrecher- und Durchfallquoten in der beruflichen Bildung aus der mangelnden Fähigkeit der Auszubildenden herrühren, (Fach-)Wissen aus Fachbüchern bzw. Lernmaterialien zu gewinnen. Ohne ausreichende Lesekompetenz mangle es Schülern, seiner Meinung nach, an einer Voraussetzung zum selbständigen Denken und Handeln (vgl. Pätzold 2010, S. 163).

Die Bedeutung von Lesekompetenz für den Berufs- und Bildungserfolg junger Menschen umfassender empirisch nachzuweisen gelang erstmals durch kanadische Daten erhoben im Rahmen von der „Youth in Transition Survey“ (YITS) von 1998. Dabei wurden junge Erwachsenen mehrmals im Rahmen einer Längsschnittstudie zu ihrem bildungs- und berufsbezogenen Werdegang befragt. Die Ergebnisse zeigen, dass fast alle der jungen Erwachsenen, die im PISA-Lesetest die Kompetenzstufen IV oder V erreicht hatten, zum Zeitpunkt der letzten Erhebung über einen High-School-Abschluss verfügten, während dies nur bei 77 Prozent der Jugendlichen, deren Leistungen der Kompetenzstufe II entsprachen, der Fall war (vgl. Naumann et al. 2010).

Es kann also auf Grund dieser Hinweise davon ausgegangen werden, dass die Lesekompetenz für den persönlichen und beruflichen Erfolg des Individuums zumindest eine nicht unerhebliche Rolle spielt. Im europäischen Rahmen wird dies auch von der Europäische Kommission (2000, S. 19) gestützt: „Die Fähigkeit, Anweisungen und Texte zu lesen und zu verstehen, ist in allen Schulfächern eine Grundbedingung für den Erfolg. Allerdings verliert die Fähigkeit, lesen und schreiben zu können, auch dann nicht an Bedeutung, wenn Kinder die Schule verlassen. Solche Fertigkeiten sind elementare Voraussetzungen für alle Bildungsbereiche ...“.

Es wird somit auch auf die über die Schulzeit hinausreichende Bedeutung des Lesens hingewiesen. Damit ist Lesen keine Fähigkeit, die ihren Raum nur in schulischen Kontexten bezüglich Qualifikationserlangung findet und abstrakt ohne konkreten Bezug zur Lebensrealität existiert. Stattdessen spielt sie einerseits unter der Perspektive des lebenslangen Lernens, also um die Leistungsfähigkeit in bildungsspezifischen Kontexten zu erhalten und zu erweitern eine Rolle, andererseits auch um konkrete tätigkeitsbezogene Aufgaben bewältigen zu können (siehe dazu Kap. 2.2).

Aebli fasst dementsprechend zusammen:

⁷ Warum und inwiefern der Beruf und ökonomischer Erfolg sowie gesellschaftliche politische Teilhabe für eine positive Lebensgestaltung relevant ist, wurde ausführlich u.a. von Gilomen (2003) dargelegt.

„Beruflicher Erfolg und berufliches Fortkommen erfordern die Fähigkeit, mit Texten umzugehen, wer es kann, kommt mit, wer darin versagt, bleibt sitzen oder geht unter, nicht nur individuell, sondern auch kollektiv, zusammen mit ganzen Wirtschaftszweigen, die im internationalen Wettbewerb abfallen“ (Aebli 1983, S. 114)

Die voran genannten Studien und erhobenen Daten beziehen sich allerdings zum Großteil auf den anglo-amerikanischen Sprachraum. Daher gilt es die Frage zu stellen, ob diese Ergebnisse auch für den Deutschland anzunehmen sind. Dahingehende großangelegte Längsschnittuntersuchungen wurden bisher nicht durchgeführt, wenn gleich es erste Erfahrungen u.a. auch aus anderen Forschungsbereichen gibt, die diese Tendenzen bestätigen (vgl. Tröster 2002; Reinberg/Hummel 2005, 2007).

Vermutungen bezüglich der Übertragbarkeit auf Deutschland werden untermauert durch die in der Schweiz (als deutschsprachiges, europäisches Nachbarland) durchgeführte TREE-Studie (Transition from Education to Employment), die besonders den Zusammenhang zwischen Lesekompetenz und Erfolg der beruflichen Erstausbildung betrachtet. Dabei zeigte sich, dass Personen mit guter Lesekompetenz mit hoher Wahrscheinlichkeit unmittelbar in die Sekundarstufe II eintreten, selten in der beruflichen Grundbildung überfordert sind, d.h. die Belastungen in Betrieb und der Berufsfachschule als eher gering empfinden, häufig einen linearen Ausbildungsverlauf vorweisen können und mit hoher Wahrscheinlichkeit einen hohen (Berufs-)Schulabschluss erreichen (vgl. Stalder 2011).

Betrachtet man als weiteren Beleg für die These, dass Lesefähigkeit, Berufserfolg und gesellschaftliches Einflusspotential nicht unabhängig sind, die Ergebnisse der 2003 u.a. in der Schweiz durchgeführten ALL-Studie (Adult Literacy and Life Skills) (vgl. Notter et al. 2006) lässt sich festhalten, dass der Anteil an guten und mittleren Lesern (Niveau 3, 4 und 5) in der deutschsprachigen Schweiz tendenziell mit jedem Einkommensquartil steigt. Dabei ist zu beobachten, dass der höchste Anteil an sehr fähigen Lesern bei den Bestverdienenden liegt. Selbstverständlich lässt sich daraus kein kausaler Zusammenhang ableiten und es gilt auch die allgemeinen kognitiven Grundfähigkeiten und Bildung als Einflussfaktoren auf beide Variablen zu berücksichtigen (vgl. Notter et al. 2006, S. 61). Wird weiterhin die Lesekompetenz im Zusammenhang mit bestimmten Berufsgruppen betrachtet, zeigt sich auch hier eine Ungleichverteilung in der Form, dass es Berufsgruppen gibt, wie die Fachkräfte in der Landwirtschaft und Fischerei oder Hilfsarbeitskräfte, in denen über zwei Drittel der Beschäftigten nicht über Niveau 2 hinauskommen, hingegen sieht es bei den Akademikern, den Techniker und den Führungskräften nahezu umgekehrt aus. Hier erreichen fast zwei Drittel mindestens Niveau 3 (vgl. Notter et al. 2006, S. 57).

Es ist anzunehmen, dass sich diese Ergebnisse auf Grund der ähnlichen Wirtschaftsstruktur (vgl. Alich 2011) mit Fokus auf dem Mittelstand und des vergleichbaren Bildungssystems auf Deutschland übertragen lassen.

Man kann als weiteren Hinweis auf die Vorteile guter Lesefähigkeiten umgekehrt auch die fünf einkommensstärksten Berufe in Deutschland betrachten. Das sind laut Statistischem Bundesamt: Geschäftsführer, Geschäftsbereichsleiter, gefolgt von Rechtsvertretern, Rechtsberatern sowie Luftverkehrsberufe, Unternehmensberater, Organisatoren und Ärzte (vgl. Statistisches Bundesamt 2006). Auch eine Betrachtung nach Wirtschaftsbereichen ist möglich – hierbei haben die folgenden Bereiche die höchsten Bruttomonatsverdienste: Finanz- und Versicherungsdienstleistungen, Energieversorgung, Information und Kommunikation, freiberufliche, wissenschaftliche und technische Dienstleistungen sowie Erziehung und Unterricht (vgl. Statistisches Bundesamt 2. Quartal 2013). Gemeinsam ist all diesen Berufen und Wirtschaftssektoren, dass von einer relativ hohen schriftbasierten Kommunikation ausgegangen werden kann, wenn gleich exakte Studien noch ausstehen. Somit ist die Annahme gerechtfertigt, dass Personen mit niedriger Lesekompetenz in diesen Berufen nur mit geringer Wahrscheinlichkeit erfolgreich sein werden und daher auf einkommensschwächere Berufe ausweichen müssen.

Als großangelegte und international vergleichende Studie unter Beteiligung Deutschlands wäre schließlich PIAAC zu nennen. Auch dort konnten Unterschiede in den grundlegenden Kompetenzen mit substanziellen Einkommensunterschieden in Verbindung gebracht werden. In allen teilnehmenden Ländern besaßen Beschäftigte mit einer höheren Lese- und alltagsmathematischen Kompetenz im Durchschnitt auch höhere Einkommen im Vergleich zu Beschäftigten mit geringeren Kompetenzniveaus. In Deutschland sind die Einkommensunterschiede zwischen Kompetenzstufe II und der höchsten Stufe mit 52 % beziehungsweise zwischen der niedrigsten Kompetenzstufe und der höchsten mit 86 % sehr deutlich. In Deutschland geht ein Zuwachs der Lesekompetenz um 50 Punkten mit knapp 10 % höherem Erwerbseinkommen pro Stunde einher (vgl. Rammstedt 2013, S. 17f).⁸

Es kann festgehalten werden:

„Es besteht ein starker Zusammenhang zwischen den grundlegenden Kompetenzen und der Partizipation am Arbeitsmarkt. In nahezu allen PIAAC-Teilnehmerländern und so auch in Deutschland besitzen 25- bis 54-jährige Erwerbstätige (das sind gemäß der internationalen ILO-Definition Personen, die mindestens eine Stunde pro Woche erwerbstätig sind) im Mittel höhere Grundkompetenzen als Erwerbslose und Nichterwerbspersonen. In Deutschland sind diese Differenzen mit 23 beziehungsweise 24 Punkten in der Lesekompetenz und 35 beziehungsweise 36 Punkten in der alltagsmathematischen Kompetenz etwas stärker ausgeprägt als im OECD-Durchschnitt“ (Rammstedt 2013, S. 16f).

In allen untersuchten Ländern zeigt sich daneben ein deutlicher Zusammenhang zwischen den Arbeitsplatzanforderungen bezüglich des Lesens und den vorhandenen diesbezüglichen Kompetenzen. Es ließ sich feststellen, dass Arbeitnehmer, die beruflich häufiger Lesetätigkeiten ausüben, durchschnittlich eine deutlich höhere Lesekompetenz aufweisen als Erwerbstätige, deren

⁸ In diesem Zusammenhang sei angemerkt, dass selbstverständlich Lesefähigkeiten nicht das alleinentscheidende Merkmal sein können, maßgebend sind auch Variablen wie Intelligenz, diverse Persönlichkeitseigenschaften und Neigungen und auch weitere Fähigkeiten bspw. mathematische Kompetenz.

Arbeitsplätze diese Fertigkeiten⁹ nie erfordern. Die Autoren räumen selbst ein, dass nicht klar ist, ob diese Zusammenhänge aus einer Selektion von Personen auf Arbeitsplätze entsprechend ihrer Kompetenzen oder aus einem (zusätzlichen) Erwerb dieser Kompetenzen auf Grund einer häufigeren Ausübung entsprechender Tätigkeiten oder aus beidem resultieren, da in PIAAC nur Querschnittsdaten erhoben wurden (vgl. Rammstedt 2013, S. 17).

Daneben lässt die Zunahme von Berufen mit hohen Leseanforderungen eine weitere Steigerung des Stellenwerts von Lesekompetenz erwarten. Insbesondere die Stabilisierung der individuellen ökonomischen Situation ist (dann) mit ungenügenden Fähigkeiten kaum möglich (vgl. Husfeldt 2001, S. 11).

Die sozialen, ökonomischen und politischen Handlungsspielräume von Personen, welche nicht über ausreichende Fähigkeiten verfügen, werden eingeschränkter, denn neben dem Einfluss auf wirtschaftliche Aspekte ist die Fähigkeit des Lesens auch essentiell, um an gesellschaftlichen Diskursen teilnehmen zu können und politischen Einfluss auszuüben, da diese vermehrt medial schriftlich (Zeitungen, online basiert) erfolgen. Denn „über die Schriftsprache werden neben Informationen und Fakten auch Ideen, Wertvorstellungen und kulturelle Inhalte vermittelt“ (Artelt et al. 2007, S. 5). Damit wird für die betroffenen Personen eine Partizipation an demokratischen Prozessen unwahrscheinlicher und erheblich erschwert.

Daneben gibt es noch den eher ideell-normativ zu verortenden Ansatz der Lesesozialisationsforschung, der die Bedeutsamkeit des Lesens aus der europäischen Bildungstradition begründet. Verkürzt dargestellt ist die Leitidee dieser Forschungsrichtung das gesellschaftlich handlungsfähige Subjekt. Bezogen auf den Leser ist dieser eine Person, die ihre Lesefähigkeit nicht nur instrumentalisiert einsetzen kann, um zu lernen, sozial aufzusteigen und Erfolg zu haben. Stattdessen soll Lesen ein „Medium der Persönlichkeitsentwicklung“ (Hurrelmann 2008, S. 22) sein, wobei Ästhetik, Moral, Empathie und kulturelle Teilhabe im Zentrum stehen (vgl. Hurrelmann 2008, S. 22f). In diesem Ansatz wird zwischen den unmittelbaren Funktionen des Lesens und den Folgefunktionen, also den aus dem Lesen herrührenden Wirkungen, die sich an den oben genannten Zielen orientieren sollten (oder zumindest daran messen lassen), unterschieden (vgl. Groeben/Hurrelmann 2004).

Neben dieser das Individuum in den Fokus stellenden Perspektive gibt es jedoch auch einige Hinweise darauf, dass die betriebs- und volkswirtschaftliche Bedeutung der Lesefähigkeit der Bürger eines Staates nicht unerheblich ist. So können bspw. kommunikative Schwierigkeiten von Mitarbeitern die Produktivität der Arbeitsprozesse im Unternehmen komplizieren (vgl. Weissenberg 2010, S. 14). Auch amerikanische Forscher sehen dies als Problem, da Industriebetriebe bereits in den späten 80er Jahren anwachsende ökonomische Probleme durch Arbeitnehmer, die nicht in der Lage sind die Leseanforderungen ihrer Arbeit zu erfüllen, bemängeln. Auch waren Bedenken betreffend der Sicherheit von Arbeitnehmern, die Warnungen nicht lesen und schriftlichen Hinweisen nicht folgen können bereits Themen in einer wachsenden Zahl von Gerichtsverhandlungen und haben zu zahlreichen Entlassungen geführt (vgl. Mikulecky 1988,

⁹ bzw. nur einen geringen Grad derselben

S. 13f). Weiterhin gibt es die Erfahrung, dass schlecht ausgebildete Arbeiter das ökonomische Problem Nummer eins sind – der Hauptfaktor der Verlangsamung des Wachstums. Es wird von über 1 Million ausgegebener US-Dollar berichtet, um die Rechtschreibfehler der Arbeitnehmer zu korrigieren und davon, dass schätzungsweise 70% der Korrespondenz korrigiert oder zumindest einmal neu geschrieben werden muss (vgl. Hymowitz 22.01.1981; Sticht/Mikulecky 1984, S. 13f). Eine Erklärung dafür lässt sich aus der besonderen Stellung der kaufmännischen Angestellten finden, denn sie „sind Träger der delegierten Funktion des Unternehmers in der Steuerung und Dokumentation industrieller Arbeitsprozesse und des Warenvertriebs einerseits und in dem Verkauf von Waren ... andererseits“ (Kaiser 2012. S. 15).

Dies wird weitergeführt in der Feststellung, dass das Individuum mehr denn je ein zentraler Bestandteil des wirtschaftlichen Erfolgs einer Nation ist und die Nation mit der anpassungsfähigsten Erwerbsbevölkerung am ehesten in der Lage sein wird ökonomische und soziale Herausforderungen zu bewältigen (Statistics Canada 1996, S. 9). Somit sei eine der besten Maßnahmen, um das neue wirtschaftliche Umfeld zu nutzen, eine Stärkung der Innovationsfähigkeit, sowie die Fähigkeit zur Anpassung an Veränderungen und Verbesserung der Produktivität seitens der Unternehmen und Arbeitsmärkte zu fördern. Allerdings hängt diese Leistungsfähigkeit primär und entscheidend von dem Wissen und Können der Bevölkerung ab. Somit ist die Beherrschung des Lesens [= der Grad der Lesekompetenz] und Schreibens ein bestimmender Faktor für Kanadas innovative und adaptive Kapazität und den zukünftigen wirtschaftlichen Wohlstand (vgl. Jones 1996, S. 39). Auch wenn Jones dies vorrangig für Kanada annimmt, gibt es wenig Grund dies nicht auch auf Deutschland zu übertragen.

Somit sollte hinreichend deutlich sein, dass die Förderung von Lesefähigkeit in der Gesellschaft und im Bildungssystem ein zentraler Punkt sein muss. Auffällig ist jedoch, dass immer wieder die Bedeutung des Lesens von unterschiedlichen Autoren logisch hergeleitet wird und überzeugend theoretisch argumentiert wird, allerdings insbesondere in Deutschland kaum belastbare empirische Studien zum Einfluss der Lesekompetenz auf eine erfolgreiche Lebensführung vorliegen.

Vor der entsprechenden Förderung stehen die Diagnose der Fähigkeiten und die genaue Kenntnis des zu diagnostizierenden Gegenstands. In diesem Zusammenhang steht diese Arbeit. Sie soll berufliches Lesen unter der besonderen Berücksichtigung der Präsentationsformate der dargebotenen Informationen untersuchen.

Dazu müssen zunächst die Leseanforderungen und -anlässe in der beruflichen Ausbildung und im Berufsalltag untersucht werden – dem widmet sich das folgende Kapitel.

2.1.2. Leseanforderungen und -anlässe in der beruflichen Ausbildung und im Berufsalltag

Nachdem auf globaler Ebene deutlich geworden ist, dass Lesekompetenz für Individuen und Gesellschaft eine wichtige Rolle spielt, bleibt die Frage, nach spezifischen Leseanforderungen

in der Arbeitswelt und für die in diesem konkreten Kontext agierenden Individuen. Unfraglich ist, dass in den meisten Berufen gelesen wird. Welche Natur, welcher Umfang, welche Häufigkeit diese Leseanlässe kennzeichnet, ist aber abhängig von der genauen Tätigkeit (vgl. Husfeldt 2001, S. 12).

Amerikanische Studien greifen dieses Forschungsfeld weitaus häufiger und intensiver auf als dies bisher in Deutschland bisher der Fall war – wobei auch hier ein deutliches Forschungsdesiderat seitens der beteiligten Forscher angemerkt wird (vgl. Diehl/Mikulecky 1980, S. 221) – deren Übertragbarkeit auf Deutschland zumindest angenommen werden kann. Daher wird an dieser Stelle auch auf sie verwiesen.

Zunächst ist einerseits festzuhalten, dass einige Forscher zu der Aussage kommen, dass die Mehrheit der Leseanlässe von Erwachsenen einen beruflichen oder arbeitsplatzbezogenen Kontext aufweist und diesbezügliches Material als Lesegrundlage dient (vgl. Mikulecky 1988, S. 1). Auch kann aufgezeigt werden, dass es einen substantiellen quantitativen Unterschied bezüglich der Leseanlässe an Arbeitstagen gegenüber nicht-Arbeitstagen gibt (vgl. White et al. 2010). Sodann gaben in einer Studie von Diehl/Mikulecky (1980)¹⁰ 61% der Befragten an, dass sie täglich bei der Arbeit mindestens eine Leseaufgabe bewältigen müssen und 22% immerhin noch, dass sie zwei bis dreimal pro Woche arbeitsbezogene Leseaufgaben ausführen müssen.

Bezüglich der Dauer des täglichen, arbeitsbezogenen Lesens gaben die Probanden im Durchschnitt 113 Minuten an, wobei es zu der paradoxen Situation kam, dass einige aussagten, nicht am Arbeitsplatz zu lesen, jedoch später (Lese-)Material aufzählten, welches sie täglich benutzen würden (vgl. Diehl/Mikulecky 1980, S. 223).

Smith kommt zu abweichenden, aber ähnlichen Ergebnissen in einer Studie¹¹ über die Lesepraxis Erwachsener: 41% der Leseanlässe fanden sich am Arbeitsplatz und 37% der täglichen Lesezeit wurde am Arbeitsplatz verbracht (ca. 54 Minuten) (vgl. Smith 2000, S. 33).

Für Deutschland existieren ähnliche Zahlen: So schreibt Fritz¹², dass sich 27% aller Lesegelegenheiten am Arbeitsplatz finden (vgl. Fritz 1991, S. 67). Problematisch an diesen Aussagen ist allerdings, dass nur nach dem Lesen von Büchern, Zeitung und Zeitschriften gefragt wurde. Wie später noch näher dargelegt wird, werden am Arbeitsplatz vermehrt Fraktale gelesen – was meint, Ausschnitte aus Text- und anderen Informationsquellen, ohne die gesamte Quelle chronologisch zu lesen. Des Weiteren lässt sich das berufliche Lesematerial kaum umfassen in diesen

¹⁰ Es wurden 107 Probanden aus 100 Berufen und 26 Arbeitsplätzen interviewt und an ihren Arbeitsstätten getestet. Die Probanden reichten von einem Rechtsanwalt und Vizepräsident eines großen Unternehmens zu Fließbandarbeitern und Steinschneidern. Sie wurden zufällig von Arbeitsplätzen in einem 70 Meile Radius von Bloomington, Indiana ausgewählt. Die Probanden repräsentierten eine breite Palette der im Wörterbuch der Berufsbezeichnungen aufgeführten Berufe und schienen repräsentativ für die erwachsene erwerbstätige Bevölkerung bezüglich Variablen wie Geschlecht, Rasse, Einkommen und Berufsgruppe zu sein.

¹¹ N=154; 88 Frauen; 84% Weiße; Angehörige der Mittelschicht; mit verschiedenen Berufen (Wirtschaft: n=40; medizinisch-pflegend: n=36; (gewerbliche + Büro-)Angestellte und Service-Mitarbeiter: n=37; Handwerker + Händler: n=35; nicht-arbeitende: n=6); Bildungshintergrund: Varianz von „ohne Schulabschluss“ (n=5) bis Doktorgrad (n=6).

¹² Die Analyse basiert auf Daten der 1988/1989 durchgeführten Studie „Kommunikationsverhalten und Medien“ (Auftraggeber: Bertelsmann Stiftung). Die repräsentative Stichprobe umfasste 3205 Bürger der BRD.

drei Kategorien abbilden, da bspw. Formulare oder produktbegleitende Texte keine Berücksichtigung finden. Somit könnte die These zu lässig sein, dass die tatsächlich mit Lesen verbrachte Zeit am Arbeitsplatz noch höher ist.

Jakobs/Schindler (2006) zitieren eine Studie von Göldi (2001), in der 134 Schweizer Ingenieure in verschiedenen Branchen zur Kommunikation am Arbeitsplatz befragt wurden. Dabei gaben etwa die Hälfte an sieben bis acht Stunden pro Tag mit Gesprächen, Lektüre oder Produktion von Texten zu verbringen, 20% gaben mindestens zwei Stunden pro Tag als Wert an und 33% noch mindestens eine Stunde (vgl. Jakobs/Schindler 2006, S. 135).

Andere Studien (vgl. Mikulecky et al. 1987; Mikulecky 1988) berichten auch von ca. zwei Stunden täglichen Lesens, was übereinstimmend zuvor auch von Sticht (1975) ermittelt werden konnte.

Verortet werden kann die mit Lesen verbrachte Zeit durch die Angabe, dass Arbeitnehmer (ausgenommen gewerbliche, wobei auch diese auf über eine Stunde kommen) durchschnittlich mehr Zeit mit Lesen verbringen als High-School-Schüler (vgl. Mikulecky/Drew 1991).

Auch ergaben verschiedene Studien (vgl. Diehl/Mikulecky 1980; Mikulecky 1988; Mikulecky/Drew 1991), dass das zu lesende Material auf oder über High-School Niveau liege und sich strukturell von diesem unterscheide, wobei die zentralen Unterschiede nach Mikulecky/Drew (1991) sind:

Arbeitsplatz	Schule/Universität
Nutzen einer Vielzahl an Materialien	Nutzen nur weniger Quelle
Hohes Level und metakognitive Fähigkeiten (Zielfestlegungen, Selbstbefragung, Beobachtung)	In erster Linie Faktenerfassung
Reading-to-do	Reading-to-learn

(vgl. Mikulecky/Drew 1991, S. 683).

Diehl/Mikulecky (1980) berichten darüber hinaus, dass am Arbeitsplatz 83% der Leseaufgaben nicht das Lesen eines ganzen Buches umfassen – im Gegenteil, meist, d.h. zu 52%, werden nur Auszüge, wie Textabschnitte oder einzelne Graphiken, im Umfang von 1-3 Seiten verwendet. Daneben zeigen sie auf, dass nur 21% der Interviewten angaben, dass Lesen entscheidend für die Fertigstellung der Arbeitsaufgabe sei, jedoch geben 56% an, dass Lesens wichtig, wenngleich nicht entscheidend für die Erfüllung der Aufgabe sei (vgl. Diehl/Mikulecky 1980, S. 223).

Lesen ist hier also nicht beschrieben als dringende Notwendigkeit, sondern als hilfreiche Fähigkeit. Die dem Text/der Graphik entnommenen Informationen hätten also auch anderweitig gewonnen werden können – beispielsweise durch Nachfragen bei Kollegen – oder die Aufgabe hätte auch ohne das Lesen erfüllt werden können – möglicherweise aber weniger gut (vgl. Diehl/Mikulecky 1980, S. 224). Beide Alternativen bergen aber die Gefahr in sich, den Arbeitsprozess weniger effizient zu gestalten, da sie entweder einen Zeitverlust oder einen Qualitätsverlust bedeuten können. Ferner ist fraglich, ob der Kollege die gewünschte Information verfügbar hat oder ebenfalls die entsprechende Quelle heranziehen würde.

Dahingestellt ist bei allen Feststellungen, wie inhärent das Lesen zu der jeweiligen Arbeitsaufgabe gehört, so dass die Vermutung legitim ist, dass noch mehr, häufiger und länger gelesen wird, nur dies den Probanden gar nicht bewusst wird (vgl. Diehl/Mikulecky 1980, S. 223f). Smith lies die Probanden zwar ein stündliches Leseprotokoll ausfüllen (vgl. Smith 2000, S. 30), jedoch ist auch hier ungewiss, ob jede Lesetätigkeit exakt protokolliert wurde oder bestimmte Situationen nicht erinnert wurden, da die eigentliche Tätigkeit im Mittelpunkt stand und so das Lesen nicht bewusst bemerkt wurde.

Wird des Weiteren das zu lesende Material und die Quellenanzahl betrachtet, so kann festgehalten werden, dass das Material verschiedener Repräsentationsmodi sein kann:

„Of the 325 reading tasks identified, 32% of the tasks involves reading textual materials only; 25% involved only the use of figures; and 37% involved reading and using both text and figures. Thus, 62% of the reading tasks involved the use of figures.”¹³ (Sticht 1977, S. 225f).

Auch White et al. (2010) konnten diesen Befund bestätigen.

Interessant an der Betrachtung des Lesematerials ist, dass dieses wenige Bezüge zur Leseintention oder -situation bietet. So kann das gleiche Material (z.B. ein Autohandbuch) von einigen Personen in einen funktionalen, arbeitsbezogenen Kontext integriert werden, von anderen in einen ebenfalls funktionalen, aber freizeitbezogenen Kontext (vgl. Kirsch/Guthrie 1977, S. 490). Ebenfalls betonen verschiedene Forscher, dass es beim arbeitsbezogenen Lesen häufig auftritt, dass die benötigten Informationen nicht nur innerhalb einer Quelle zu finden sind, sondern mehrere Quellen herangezogen werden müssen (vgl. Mikulecky 1988, S. 5; Mikulecky/Drew 1991, S. 671)

Relevant und aufschlussreich scheint die Betrachtung der Leseintention der Probanden zu sein. Alle Studien stimmen überein, dass die Kernmotivation des Lesens darin liegt, eine arbeitsbezogene Handlung auszuführen (vgl. Sticht 1977; Diehl/Mikulecky 1980; Mikulecky et al. 1987; Mikulecky 1988; Mikulecky/Drew 1991). So geht bspw. aus Stichts (1977) Studie¹⁴ hervor,

¹³ Ungefähr: „Von den 325 identifizierten Leseaufgaben bezogen 32% der Aufgaben das Lesen von „nur-Text-Material“ ein; 25% bezogen die Benutzung von „nur-Schaubildern“ ein und 37% bezogen Lesen und Benutzen von beidem, Text und Schaubild, ein. Demnach bezogen 62% der Leseaufgaben Schaubilder ein.“

¹⁴ Die Daten wurden durch ein Interview (Dauer ca. 1 - 1 1/2h) mit 178 Navyangehörigen, die sich in drei verschiedenen Stadien des „Ausbildungsprogramms“ (Lernender, Ausführender, Lehrender) befanden. Frauen machten nur 3% der Stichprobe aus, 81% waren Europäer.

dass 60% der Leseaufgaben Aufgaben im Handlungskontext einer spezifischen Arbeitssituation waren (bezeichnet als: reading-to-do). Auch Diehl/Mikulecky (1980) kommen zu ähnlichen Ergebnissen (63%). In ihrer Studie wird nur in 11% der Fälle Lernen als Leseziel (bezeichnet als: reading-to-learn) angegeben. Auch Smith (2000) ermittelt, dass 15% des täglichen Lesematerials aus funktionalen Gründen (bspw. Informationen gewinnen, um eine Aufgabe auszuführen) gelesen wird und die Leseanlässe zu 34% ihren Grund in der Erfüllung einer Arbeitsaufgabe hatten (vgl. Smith 2000, S. 33f; 36).

Reading-to-learn oder epistemisches Lesen ist gekennzeichnet durch die Intention des Subjekts die Textinformationen zu behalten – also zu „lernen“. Dabei kann auch oftmals der Einsatz von Lernstrategien beobachtet werden. Reading-to-do ist im Gegensatz dazu gekennzeichnet durch die Nutzung des Materials primär als Hilfestellung zu einer Handlung (bspw. der Reparatur einer Maschine) und später nicht oder nur zufällig über das Behalten der Informationen berichtet werden kann. Hierbei dient das Material also als externer Gedächtnisspeicher. Reading-to-do Leseaufgaben sind daher meist in eine andere, konkrete berufliche Aufgabe integriert, die gewonnenen Informationen werden unverzüglich auf diese Aufgabe angewandt und es ist bemerkenswert, dass sich die Leseaufgaben immer wieder wiederholen können, ohne dass es unmittelbar zu einem Lerneffekt kommt. Sauer (1999) verweist dazu auf Pander Maat, welcher in den 90er Jahren Charakteristika beruflicher Leseaufgaben (für die Niederländische Sprache) ermittelt hat und diese kennzeichnet durch einen engen Zusammenhang zwischen dem Ergebnis der Textverarbeitung und den weiteren Tätigkeiten des Lesers. Diese Art der Beziehung separiert den eigentlichen Leseprozess von der Tätigkeit, die das Ergebnis der Textverarbeitung – also mindestens den Aufbau einer mentalen Repräsentation des Textes – benötigt, um ausgeführt werden zu können. Der Leseprozess wird damit quasi zu einer Aufgabe in einer komplexen Handlung verschachtelt – in dieser Handlung können auch mehrere separate Leseprozesse stattfinden, wobei der eine den anderen bedingt und so fort. Die anschließende Handlung kann nur ausgeführt werden, wenn das vorherige Lesen zu umsetzbaren Ergebnissen geführt hat (vgl. Sauer 1999, S. 155f) oder falls dies nicht der Fall ist alternative Quellen akquiriert werden können. Daneben hat hierbei der Leser zumeist Zugang zu außersprachlichen Hinweisen, die für gewöhnlich einen direkten Bezug zum Lesematerial haben – oft existiert sogar eine 1:1 Bezugnahme zwischen Lesematerial und Arbeitsumgebung (vgl. Diehl/Mikulecky 1980, S. 223, 225f).

Die zunehmende Bedeutung dieser Forschung wird gestützt durch den damals bereits getätigten Ausblick, dass wahrscheinlich nicht für alle neuen Arbeitsplätze Jahre der Ausbildung erforderlich sein werden, jedoch anzunehmen ist, dass höhere Mindestansprüche an basale Grundkenntnisse aufkommen werden. Die Aufgaben eines Hausmeisters umfassen bereits heute mehr, als nur einfache manuelle Tätigkeiten und vermutlich wird es erforderlich sein, noch mehr zu können, da zukünftig die Verantwortung für die komplexe technische Ausrüstung der Gebäude zumindest zum Teil übernommen wird. Berufe, die keine oder nur geringe Grundkenntnisse erfordern, werden zunehmend schneller verschwinden, während die neu geschaffenen Berufe von den Arbeitnehmern Lesen, Schreiben und Rechnen auf einem verhältnismäßig

hohen Kompetenzniveau erfordern werden, um die täglichen Probleme der Arbeitswelt zu lösen (vgl. Sticht/Mikulecky 1984, S. 5).

Sticht (1977) hält in diesem Kontext für besonders erwähnenswert, dass die meiste ihm bekannte Forschung sich mit Lernen aus Texten befasst. Die vorliegenden Ergebnisse und die Metaanalyse der vorgestellten Studien legen jedoch den Bedarf und die Notwendigkeit weitere Forschung bezüglich der Nutzung von Texten zur Ausführung einer Handlung/im Handlungskontext nahe (vgl. Sticht 1977, S. 225). Darin stimmen die amerikanischen Forscher insgesamt überein und konstatieren trotz der Bedeutung dieses Feldes massive Lücken. Dazu gehört für sie vorrangig die Entwicklung eines „Arbeitsplatzlesefähigkeitsprozessmodells“ (Mikulecky/Drew 1991, S. 683), die Erforschung der Eigenschaften des Leseprozesses am Arbeitsplatz, die Nutzung von Texten (und Graphiken) zur Bewältigung einer Handlung (im Unterschied zum Lernen aus Texten) oder auch welche Lesefähigkeiten eine Person konkret haben muss, um in typischen Arbeitssituationen kompetent handeln zu können (vgl. Sticht 1977; Diehl/Mikulecky 1980; Mikulecky/Drew 1991).

Umfangreiche Studien zu Quantität und Qualität beruflicher Leseanlässe in Deutschland stehen bisher noch aus, wenngleich der Stellenwert dieses Feldes offenkundig geworden sein dürfte und einige Projekte zur Erfüllung dieses Forschungsdesiderats angelaufen sind.

So kann bisher zumindest unstrittig festgehalten werden, dass Auszubildende als angehende Fachkräfte in ihrer Ausbildung und später im Beruf fachbezogene Inhalte aus Fachbüchern, Fachberichten und Produktspezifikationen erschließen müssen. Sie müssen betriebliche Informationssysteme von der Leistungsbeschreibung bis hin zur Dokumentation der erbrachten Leistungen zu nutzen wissen (vgl. Becker-Mrotzek et al. 2006, S. 15). Wie und in welchem Umfang dies jedoch konkret für die einzelnen Berufe oder Berufsfelder aussieht ist bisher wenig erforscht.

Zu betrachten wäre dabei zunächst, ob und in welchen Anforderungssituationen Auszubildende tatsächlich in der betrieblichen Realität lesen müssen und die Frage, welchen Stellenwert (funktionales) Lesen jeweils einnimmt (vgl. Ziegler/Gschwendtner 2010).

Dieses Desiderat füllen bisher nur wenige Arbeiten explizit für den Bereich der beruflichen Leseanforderungen, wenngleich sich andere Arbeiten allgemein mit sprachlich-kommunikativen Anforderungen am Arbeitsplatz befassen. Dazu muss allerdings zunächst auch festgehalten werden, dass der Anspruch, Leseanforderungen im beruflichen Kontext zu erfassen, in Anbetracht der Variationen beruflicher Lesekontexte äußerst ambitioniert ist. Allein die domänenspezifischen Differenzen zwischen kaufmännisch-verwaltenden und gewerblichen oder pflegerischen Tätigkeiten sind vermutlich erheblich. Daneben beeinflussen branchenspezifische, betriebliche und arbeitsorganisationsbezogene Faktoren die Leseanforderungen im Berufsfeld entscheidend (vgl. Ziegler et al. 2012, S. 4).

Gegenwärtig wären da besonders die Arbeiten von Keimes/Rexing hervorzuheben. Es wurden exemplarisch die Ausbildungsberufe Maurer/in bzw. Straßenbauer/in untersucht, die charakterisiert werden können durch eine Schülerklientel mit eher niedrigen kognitiven Eingangsvoraussetzungen und teilweise erheblichen Defiziten im Bereich der Lesekompetenz. Zur korrekten Ermittlung aller Leseanlässe ist eine gründliche fachwissenschaftliche Analyse der im Lehrplan aufgeführten beruflichen Tätigkeiten notwendig (vgl. Keimes/Rexing 2011b, S. 2f). Es wurde zunächst eine Dokumentenanalyse der entsprechenden Richtlinien/Lehrpläne bzw. Ausbildungsordnungen durchgeführt und als Ergebnis der Analyse konnte festgehalten werden, dass sich in den Curricula zur beruflichen Grundbildung nur drei explizite Nennungen von Leseanlässen finden lassen (vgl. Anhang Tabelle 42). Bemerkenswert ist dabei weiterhin, dass sich diese ausschließlich auf depiktionale Strukturen beziehen (z.B. Zeichnungen, Pläne). Darüber hinaus konnten sie jedoch diverse implizite Leseanlässe identifizieren (z.B. Bestimmen der Zusammensetzung des Betons anhand von Tabellen). Für die Fachstufen kommen sie zu vergleichbaren Ergebnissen über die Quantität von Leseanlässen im Lehrplan des Bildungsgangs Maurer/in. Bezogen auf die verbleibenden Ausbildungsjahre finden sich drei explizite Nennungen von Leseanlässen, darüber hinaus aber in allen Lernfeldern zahlreiche weitere implizite Nennungen (vgl. Keimes/Rexing 2011b, S. 4)

Bei der Erforschung der ausbildungsrelevanten Textsorten von angehenden Land- und Baumaschinenmechanikern kommen Keimes et al. zu vergleichbaren Erkenntnissen. Genannt wird vorrangig das Werkstatthandbuch als wichtigste Quelle, dann folgend Fachzeitschriften, Tabellen, Betriebsanleitungen und Diagramme als zentral ausbildungsrelevante Textsorten. Als konkrete Leseanlässe innerhalb der beruflichen Ausbildung werden von den Auszubildenden am häufigsten der Berufsschulunterricht und die Prüfungsvorbereitungen genannt. Weiterhin sind nach Angaben der Auszubildenden auch das Nachschlagen in Handbüchern im Zusammenhang mit Montagen und Reparaturen, bei Inventuren, Inspektionen und Ölwechseln exemplarische Lesesituationen. Besonders auffällig für Keimes et al. war im Kontext der Befragung die Schwierigkeiten der Schüler, spontan Textmaterial bzw. Leseanlässe aus ihrem beruflichen Alltag zu nennen. Es mussten mehrfach Impulse gegeben und wiederholt nachgefragt werden, bis die Schüler entsprechende Antworten äußerten (vgl. Keimes et al. 2011, S. 228).

In diesem Bereich besonders hervorzuheben, da sehr umfassend, ist die Arbeit von Keimes: „Lesen. Eine empirische Studie zur berufsspezifischen Relevanz des Lesens in gewerblich-technischen Bildungsgängen“ (2014). Dort werden vertiefte Untersuchungen zum Berufsfeld Bautechnik in den Berufen Maurer und Straßenbauer angestellt. Bezüglich der Stichprobe und dem methodischen Vorgehen ist festzuhalten, dass einerseits leitfadengestützte Experteninterviews mit 14 Akteuren der beruflichen Bildung, die über einschlägige Erfahrung in der Baustellenpraxis verfügten und sieben leitfadengestützte Gruppeninterviews mit Auszubildenden zum Maurer bzw. Straßenbauer (N=32), die sich zum Zeitpunkt der Befragung in der Mittel- oder Oberstufe befanden, durchgeführt wurden. Das gewonnene Material wurde dann einer Qualitativen Datenanalyse sowie einer Qualitativen Inhaltsanalyse nach Mayring unterzogen.

Die zentralen Ergebnisse bezogen auf Auszubildende sind (stark verkürzt dargestellt):

„Lesefähigkeit scheint aus Sicht der Auszubildenden für die Aufnahme einer Berufsausbildung zum/zur Maurer/in bzw. Straßenbauer/in ... eine eher subordinierende Bedeutung zu haben und wurde insgesamt lediglich von zwei Personen genannt.“ (Keimes 2014, S. 147)

Es „dominieren mündliche Kommunikationsstrukturen im Baustellenbetrieb. Überwiegend sind es Personen mit Führungsverantwortung (z.B. Meister und Poliere), die Arbeitsanweisungen in schriftlicher Form erhalten.“ (Keimes 2014, S. 148)

Es gibt nur „wenige berufliche Anforderungssituationen, in denen Auszubildende Informationen aus schriftlichen Dokumenten entnehmen müssen.“ (Keimes 2014, S. 150)

Die „Leseaktivitäten [beziehen sich] nahezu ausschließlich auf das Lesen von Zeichnungen/Plänen“ (Keimes 2014, S. 150)

Die „Lesehäufigkeit bei der Bewältigung beruflicher Anforderungssituationen (im betrieblichen Kontext) [werden] ebenfalls als eher gering eingeschätzt.“ (Keimes 2014, S. 151)

Verglichen werden diese Ergebnisse in einem nächsten Schritt mit den Anforderungen, die an die Gesellen/Facharbeiter gestellt werden, da dies die Aufgaben sind, die die Auszubildenden zukünftig erfüllen sollen. Dabei zeigen sich große Diskrepanzen zu den Anforderungen an Auszubildende. Bei der Frage nach spezifische Charakteristika eines qualifizierten Gesellen/Facharbeiters werden mit größerer Häufigkeit auch Lesen, Verstehen und Umsetzen von Plänen angegeben, obwohl diese in der Ausbildung keine so große Rolle spielen.

Dass mit zunehmender Führungsverantwortung auch die Leseanforderungen steigen, wurde an Hand einer inhaltsanalytischen Auswertung der (ausbildungs-)relevanten Texte gezeigt. Diese wurden zum Zwecke einer direkten Vergleichbarkeit tabellarisch und differenziert für die Personengruppen Auszubilden, Gesellen/Facharbeiter und Personen mit Führungsverantwortung (z.B. Meister, Poliere) erfasst (vgl. Anhang Tabelle 43). Dabei konnte festgehalten werden, dass für Auszubildende nur schriftliche Dokumente als ausbildungsrelevant bezeichnet werden, dabei dominieren in quantitativer Hinsicht die Zeichnungen deutlich:

„Die Auszählung des Begriffs Zeichnung bzw. entsprechender Synonyme (Plan, technische Skizze) ergab, dass diese seitens der Befragten über 100 Mal genannt wurden, während Arbeitsanweisungen (25 Nennungen), Produkt- und Verarbeitungshinweise (17 Nennungen) oder Leistungsverzeichnisse (16 Nennungen) deutlich seltener erwähnt wurden (14)“ (Keimes 2014, S. 155).

Daraus wird der Schluss gezogen, dass davon ausgegangen werden kann, dass besonders Bauzeichnungen für die betriebliche Praxis eine herausgehobene Bedeutung zukommt (vgl. Keimes 2014, S. 155).

Im Gegensatz dazu ist die Bandbreite an Texten für Gesellen breiter und beinhaltet neben den ausbildungsrelevanten Texten auch Lieferscheine, gesetzliche Vorschriften und Normtexte. Dass dann diese Texte nicht auch von Auszubildenden gelesen werden müssen, kann als überraschend markiert werden. Überzeugend wird an dieser Stelle argumentiert, dass angesichts

des Ziels der Berufsausbildung – auf eine sich anschließende Berufstätigkeit vorzubereiten und den Erwerb der erforderlichen Berufserfahrung zu ermöglichen – eine Heranführung an die aufgeführten Texte gewissermaßen naheliegend wäre (vgl. Keimes 2014, S. 155).

Abschließend wurden Personen mit Führungsverantwortung untersucht. Diese übernehmen nach Keimes im Baustellenbetrieb neben Bauarbeiten auch planerische, organisatorische, kaufmännische und verwaltende Aufgaben. Entsprechend werden für Personen mit bauleitender Funktion solche Texte in den Interviews berichtet, die z.B. in den Bereich administrativer Tätigkeiten (betriebsbezogene E-Mails/Schriftverkehr) und der Kommunikation mit externen Akteuren (z.B. Genehmigungen) eingeordnet werden können (vgl. Keimes 2014, S. 155).

Zusammenfassend kommt sie zu der Beobachtung, dass Leseaufgaben hauptsächlich Personen mit Führungsverantwortung und in nur geringerem Maße Gesellen/Facharbeitern obliegen. Schlussfolgernd kann so allerdings Lesekompetenz im Umkehrschluss unter Aspekten der Berufskarriere als ein zentrales Instrument qualifiziert werden (vgl. Keimes 2014, S. 155f).

Eine besonders interessante Erkenntnis ist, dass Lesen zwar einerseits zweifellos als wichtige Fähigkeit evaluiert, andererseits aber in Diskrepanz hierzu im Kontext der betrieblichen (Ausbildungs-)Praxis nicht vermittelt wird¹⁵. Somit lautet Keimes These: „Im Hinblick auf die Relevanz des Lesens artikuliert sich ein fragiles Verhältnis zwischen Normativität und betrieblicher Realität.“ (Keimes 2014, S. 161). Es werden verschiedene Hinweise darauf gefunden, dass die Äußerungen zur Bedeutung des Lesens mitunter inkonsistent und paradox sind, da bisweilen gravierende Abweichungen zwischen expliziten und impliziten Bedeutungszuschreibungen bestehen. Dass es sich dabei seitens der Befragten möglicherweise auch um Antworten im Sinne sozialer Erwünschtheit handelt, ist zwar nicht auszuschließen (vgl. Keimes 2014, S. 161), dennoch scheint diese Beobachtung konsistent zu der in Kapitel 2.1.1 gelangten Erkenntnis.

Als weiteres Forschungsergebnis zeigt Keimes die textrezeptiven Handlungsfelder auf (vgl. Tabelle 44). Diese werden mit den schulischen Lernfeldern verglichen, wobei sich die Relation der textrezeptiven Handlungsfelder und der schulischen Lernfelder bzw. der zugrunde liegenden beruflichen Handlungssituationen zeigt. Die textrezeptiven Handlungsfelder schneiden die schulischen Lernfelder sozusagen quer, sie liegen gewissermaßen horizontal zur Struktur der Lernfelder (vgl. Keimes 2014, S. 174).

Abschließend betrachtet Keimes die Repräsentationsformate der beruflichen Texte und kommt zu der Beobachtung, dass es sich bei den meisten Texten um sogenannte gemischte Formate handelt, die gleichermaßen Merkmale deskriptionaler als auch depiktionaler Repräsentationen aufweisen (vgl. Tabelle 45).

Ferner wird in der Arbeit erkennbar, dass das Lesen der meisten Texte bzw. der Bewältigung entsprechender betrieblicher Anforderungssituationen zumindest integrierende, in vielen Fällen sogar generierende Leistungen im Hinblick auf die Lesekompetenz erfordert. Das heißt, die

¹⁵Ähnliche Beobachtungen machten Baumann und Siemon 2013 in ihren Untersuchungen bezogen auf notwendige Schreibkompetenzen von Auszubildenden (Baumann und Siemon 2013).

Bewältigung textrezeptiver Anforderungen setzt in diesen Fällen die Konstruktion eines mentalen Modells voraus und kann als komplex bezeichnet werden (vgl. Keimes 2014, S. 176f).

Eine ähnliche Studie, wenn auch mit leicht anderer Fragestellung, von Kaiser (2012) bezüglich der „Gemeinsamkeiten und Unterschiede kaufmännisch-betriebswirtschaftlicher Ausbildungs- und Fortbildungsberufe (GUK)“ kommt u.a. zu der Erkenntnis, dass kaufmännische Tätigkeiten geprägt sind durch die Vermittlung zwischen realen Prozessabläufen und deren Abbildungen in Rechnungen, Belegen, Grafiken und Schriftsätzen, sowie der Verarbeitung und Transformation von Information (vgl. Kaiser 2012, S. 14). Dabei wird vornehmlich mit symbolischen Repräsentationen operiert, die der Decodierung bedürfen, also müssen sie für den Kernbereich ihres Geschäfts über gute Lesekompetenz verfügen.

Eine eigene Analyse der Curricula ausgewählter Berufe, die ein breites, dennoch nicht repräsentatives Spektrum an Berufen umfasst, stützt diese Erkenntnis. Untersucht wurden die Berufe „Elektroniker/in für Automatisierungstechnik“, „Elektroniker/in für Betriebstechnik“, „Industriekaufleute“, „KFZ-Mechatroniker/in“, „Kaufleute für Spedition und Logistik-dienstleistungen“, „Kaufleute EFZ“ (Schweizer Ausbildungsberuf), „medizinische/r Fachangestellte/r“ und „Altenpfleger/in“.

Zentrale Beobachtungen sind in Tabelle 1 festgehalten.

Tabelle 1 – exemplarische Leseanlässe der beruflichen Ausbildung gemäß Lehrplan (eigene Darstellung)¹⁶

Kategorie	Elektroniker/in für Automatisierungstechnik	Elektroniker/in für Betriebstechnik	KFZ-Mechatroniker/in	medizinische/r Fachangestellte/r	Kaufleute für Spedition und Logistikdienstleistungen	Industriekaufleute	Kaufleute EFZ	Altenpfleger/in
	Deskriptionale Repräsentationsformate							
Darlegung/Beschreibung	<u>Umgang</u> mit technischen Dokumentationen, <u>Erstellen</u> von Fehler- und Schwachstellenanalysen	<u>Verwendung</u> technischer Dokumentationen	<u>Erstellen</u> und <u>Verwenden</u> technischer Dokumentationen	<u>Führung</u> der Ablage und Archivierung	Gefahrgutklassifikationen und <u>Umgang</u> mit Gefahrgut	ordnungspolitischer Rahmen von Unternehmen, Strukturpolitik, staatliche Ordnungspolitik	<u>informieren</u> über Aufgaben, Arbeitsanforderungen, Tätigkeiten und exemplarische Arbeitsprozesse des Berufes	<u>Befassen</u> mit neuen pflege- & sozialwissenschaftlichen Erkenntnissen
Anweisung	Normen, Betriebsanleitungen	Betriebsanleitungen	Betriebsanleitungen, Haftungsrecht	Betriebsanleitungen, Haftungsrecht	Vertragsrecht, Haftungsrecht	Gesetzliche Vorschriften, Vertragsrecht, Haftungsrecht	Vertragsrecht	Haftungsrecht, Gesetze zur Sorgfalts- und Aufsichtspflicht
	Hybridformen							
Hypertexte	Onlinehilfe	rechnergestützte Informationsbeschaffung/Informationsgewinnung	rechnergestützte Informationsbeschaffung/Informationsgewinnung	rechnergestützte Informationsbeschaffung/Informationsgewinnung	rechnergestützte Informationsbeschaffung/Informationsgewinnung	rechnergestützte Informationsbeschaffung/Informationsgewinnung	rechnergestützte Informationsbeschaffung/Informationsgewinnung	

¹⁶ Quellen: Auslegung durch die Berufsschulen in Hessen 2008; Beschluss der Kultusministerkonferenz 2001; Beschluss der Kultusministerkonferenz 2003a; Beschluss der Kultusministerkonferenz 2003b; Beschluss der Kultusministerkonferenz 2003c; Beschluss der Kultusministerkonferenz 2005; Beschluss der Kultusministerkonferenz 2004; Hessischen Sozialministerium 2011; Industrie- und Handelskammer Frankfurt am Main 2004; Ministerium für Arbeit, Gesundheit und Soziales des Landes Nordrhein-Westfalen 2006; Ministerium für Schule und Weiterbildung 2010; Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen 2003; Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen 2004b; Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen 2004c; Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen 2004a

		mit modernen Medien	mit modernen Medien	mit modernen Medien	mit modernen Medien	mit modernen Medien	mit modernen Medien	
Dokumente oder Register	<u>Interpretation</u> von Funktions- und Fehlerprotokollen, <u>Erstellen</u> von Inbetriebnahmeprotokollen	<u>Erstellen</u> technischer Dokumentationen	Servicekonzepte und -umfänge; technische Informations-, Kommunikations-, Dokumentationssysteme	<u>Verwaltung/Erstellung</u> von Patientendaten /Patientenkartei	Begleitpapiere, Ausfuhrdokumente, Spedititeurversanddokumente	<u>Bedarfsermittlung</u> auf Basis vorgegebener Stücklisten bzw. Leistungsmerkmale	<u>Planung</u> von Auftragsabwicklung nach Kundenanforderungen	Pflegedokumentation
Formulare	<u>Analyse</u> von Aufträgen	Servicepläne/Wartungspläne		Vorbereiten und <u>Ausfüllen</u> von Formularen	Beantragung von Zollverfahren	Bestellvorschläge unter Berücksichtigung der Wiederbeschaffungszeiten und Verbrauchsschätzungen <u>erarbeiten</u>	Bestellabwicklung	bei der <u>Erstellung</u> von Anträgen helfen
Tabellen und Matrizen	tabellarische und grafische Messdatenpräsentation	tabellarische und grafische Messdatenpräsentation	tabellarische und grafische Messdaten <u>präsentation</u>	Erkennen von Auffälligkeiten bei Untersuchungsergebnissen	Lagerverwaltung, <u>Ermittlung</u> des Bedarfs an Sachgütern des eigenen Unternehmens	<u>Bedarfsermittlung</u> auf Basis vorgegebener Stücklisten bzw. Leistungsmerkmale	Tabellenkalkulation, <u>Vergleich</u> der Produktpaletten versch. Unternehmen	<u>Umgang</u> mit Bestands- & Verbrauchslisten

Aufrufe und Anzeigen	Angebotserstellung	Angebotserstellung	Pflege der Kunden-Lieferanten-Beziehung	Zusatzuntersuchungen, Serviceleistungen	Pflege der Kunden-Lieferanten-Beziehung	Angebotserstellung	Maßnahmen des Marketings	
Verträge	<u>Mitwirken</u> bei Verträgen			Behandlungsvertrag	Speditionsverträge	<u>Anbahnen</u> und <u>Abschluss</u> von Verträgen	<u>Mitwirken</u> bei Verträgen	Pflegevertrag
Informationsblätter	<u>Erstellen</u> von Fehler- und Schwachstellenanalysen	<u>Erläuterung</u> von Angeboten und Kosten gegenüber dem Kunden	Reparaturleitfäden	<u>Kundenberatung/Patientenberatung</u> , Information über Anwendung, Neben- und (Wechsel-) Wirkung sowie Risiken	Begleitpapiere, Ausfuhrdokumente, Speditieurversanddokumente	(Personal-) Statistiken	Kundenberatung/Patientenberatung	Bedienungsanleitung med. Ausrüstung
	Depiktionale Repräsentationsformate							
Schaubilder/Diagramme	Prozessvisualisierung	<u>Erstellen</u> von Schalt- und Installationsplänen, Nutzung von Funktionsschemata	Schaltpläne, Schaltzeichnungen	Planung der Pflege und Wartung von Instrumenten und Geräten	<u>Erstellung</u> von Tourenplänen zur Durchführung des Distributionsprozesses	Standortfaktoren, <u>Verwendung</u> und <u>Erstellung</u> von Diagrammen	Verständnis der Sozialen Marktwirtschaft, Modell des Wirtschaftskreislaufs	Dienstpläne

Hervorzuheben ist, dass eine Systematisierung durch eine Einteilung in die Kategorien „depiktionale Repräsentationen“, „deskriptionale Repräsentationen“ und „Hybridformen“ erfolgreich vollzogen werden kann (deren Merkmale werden ausführlich in Kapitel 3.1.3 erläutert). Diese Kategorien werden nochmals unterteilt in die Sub-Kategorien:

- Darlegung/Beschreibung
- Anweisung
- Dokumente oder Register
- Hypertexte
- Formulare
- Tabellen und Matrizen
- Informationsblätter
- Aufrufe und Anzeigen
- Bescheinigungen
- Schaubilder/Diagramme

Schon an den Subkategorien ist die Dominanz von Hybridformen erkennbar. Auch wird deutlich, dass selbst wenn nicht jede Subkategorie für jeden Beruf ein Pendant findet, doch die meisten Subkategorien vertreten sind und in jeden Fall die Kategorien depiktionale Repräsentationsformate/Hybridform/deskriptionale Repräsentationsformate aufgegriffen werden.

Interessant ist, dass – wie auch bei Keimes Analysen – beobachtet werden kann, dass nur selten wirklich von „Lesen“ gesprochen wird, meist wird eher die textproduktive Seite (erstellen, führen, ausfüllen, mitwirken → in Tabelle 1 grün unterstrichen) betont, die jedoch erfahrungsgemäß vorher eine textrezeptive Handlung voraussetzt.

Die zweite wichtige Erkenntnis stützt die Analysen der amerikanischen Forscher bezüglich des Konstruktes „reading-to-do“. Lesen wird immer in bestimmte berufliche Handlungskontexte integriert (Umgang mit..., Verwendung von..., informieren über..., befassen mit..., Interpretation von..., Planung von..., Ermittlung von..., Vergleich von..., Erläuterung..., Beratung... → in Tabelle 1 orange unterstrichen). Das zeigt, dass Lesen hier nicht das primäre Ziel/die primäre Aufgabe ist, sondern Mittel (Werkzeug), um etwas anderes zu erreichen. Damit wird allerdings wiederum belegt, dass Lesen immens wichtig ist, um in der beruflichen Ausbildung Erfolg zu haben, da wie bereits ausgeführt diese verschiedenen Handlungsziele wahrscheinlich auch mehr oder weniger ohne die Nutzung von schriftlichem Material erreicht werden könnten, dann jedoch die Effizienz bezüglich Zeit- und Personalressourcen beeinträchtigt wäre – ganz davon abgesehen, dass angenommen werden kann, dass es zu erheblichen Störungen im Arbeitsablauf käme, sollte ein Auszubildender in kurzen regelmäßigen Abständen mit Nachfragen an Kollegen herantreten, welche eigentlich durch minimale eigene Recherchen beantwortet werden könnten.

Um weiterhin die Aussage, dass die Leseanforderungen am Arbeitsplatz nicht unerheblich sind, zu stützen, wird im Folgenden allgemeiner auf sprachlich-kommunikative Handlungsfelder rekurriert. Wenngleich diese nicht explizit von Lesen sprechen, beziehen sie doch auch oft schriftliche Repräsentationsformen bezüglich der Informationen ein.

Brünner (1998) hält bezogen auf betriebliche Kommunikation fest, dass es in Unternehmen bestimmte Schwerpunkte sprach- und kommunikationsintensiver Tätigkeiten gibt, wie (planende, koordinierende, kontrollierende) Management-Aufgaben, Aus- und Fortbildung oder Aktivitäten der technischen Dokumentation (Speicherung und Weitergabe technischen Wissens) sowie betriebsexterne Kommunikation (Werbung, Verkauf, Service und Reklamationsabwicklung), welche sich vorwiegend auf die Außenbeziehungen des Unternehmens zu den Kunden richten. Mit der Funktionalisierung eben dieser Kommunikation geht ihre Formalisierung einher, d.h. erhebliche Anteile des kommunikativen Handelns sind beständig und personenunabhängig geregelt – z.B. im Hinblick auf Adressaten, Inhalt, Form, Medium, Zeitpunkt oder Umfang. Ferner ist eine Institutionalisierung und Verrechtlichung der Kommunikation zu beobachten und es ist offenkundig, dass Kommunikation für Unternehmen einen Kostenfaktor darstellt und deshalb Ökonomieprinzipien unterworfen wird. Solche Ökonomieprinzipien werden etwa in der Formalisierung der Kommunikation und im Einsatz technischer Medien verfolgt (vgl. Brünner 1998, S. 635).

Auch wenn hierbei konkret weder Leseanlässe noch Lesen explizit genannte werden, ist offensichtlich, dass technische Dokumentation, Planung oder Werbung nicht auf einer rein medial mündlichen Ebene ablaufen können. Dauerhaftigkeit und Personenunabhängigkeit sind Merkmale schriftbasierter Kommunikation. Eine dementsprechende Formalisierung mündlicher Kommunikation in großem Umfang ist unter zeitökonomischen Aspekten nur schwer vorstellbar, wenngleich sie in sprachlichen Ritualen stattgefunden hat.

Jakobs (2008a) nennt allgemein einige Gründe für die wachsende Anzahl an schriftlich (text-rezeptiv oder -produktiv) zu bearbeitenden Aufgaben, welche in diesem Kontext erhellend sein können. Dazu gehört einerseits die Aufwertung von Wissen als Ressource, d.h. Information und der Zugriff auf sie gelten vermehrt als wertschöpfende Ressourcen (siehe auch: die Entwicklung der sogenannten Wissens- oder Informationsgesellschaft) und ohne den Zugriff auf neues – wie vorhandenes – Wissen, sind weder Innovationen noch die Optimierung von Produktionsprozessen möglich. Schriftbasierte Formate gelten nach wie vor als wichtigstes Mittel der Fixierung, Kommunikation und Tradierung von Wissen. Des Weiteren kann als ein weiterer Grund die Standardisierung von Arbeit, was meint, dass Arbeitsaufgaben, -prozesse und -ergebnisse zunehmend vereinheitlicht werden und ebendiese Normen in Dokumenten fixiert werden. Ferner ist der steigende Wunsch nach Professionalisierung von Berufen als Grund zu nennen, denn schriftliche Erfassung von Arbeitsprozessen und -produkten macht Leistung vergleich- und bewertbar. Insgesamt tragen Regulierung und Vergleichbarkeit von Leistung zu dieser Professionalisierung bei und damit oft auch zu ihrer Aufwertung. Des Weiteren befördert der inflationäre Einsatz elektronischer Informations- und Kommunikationsmittel wie PC, Notebook, Smart-

phone, PDA, Internet und Intranet die Renaissance des Schreibens, wenngleich in transformierter Form, denn Arbeitsaufgaben, die früher primär mündlich kommuniziert wurden (wie Anweisungen, Absprachen, Vereinbarungen etc.), werden zunehmend schriftlich transportiert (SMS statt Anruf, E-Mail statt Meeting). Weiterhin vermutet Jakob die Industrialisierung der Kommunikationsarbeit als einen weiteren Grund für die wachsenden Anforderungen an die schriftsprachlichen Fähigkeiten der Arbeitnehmer. Darunter versteht sie, dass die Elektronisierung von Kommunikationsarbeit, welche eine wesentliche Voraussetzung für die schrittweise Automatisierung von Kommunikation darstellt und die Modularisierung und Standardisierung kommunikativer Prozesse befördere. Abschließend konstatiert Jakobs, dass die notwendig gewordene rechtliche Absicherung, sowie der rechtliche Absicherungsdruck zunehme und so der Anteil schriftlicher Kommunikation steigen müsse, um das Handeln von Personen und Unternehmen dokumentieren und ratifizieren zu können (vgl. Jakobs 2007, S. 27f, 2008a, S. 2, 2008b, S. 257).

Jakobs hat neben diesen theoretischen Betrachtungen an Hand einer kleinen Stichprobe aus verschiedenen Berufen¹⁷ Textproduktivität am Arbeitsplatz hinsichtlich Umfang und Bedeutung untersucht. Dabei kommt sie zu der Erkenntnis, dass insgesamt vieles schriftlich erfasst wird und Schreiben in vielen Berufen und Arbeitsumgebungen als wichtiger Bestandteil der Tätigkeit betrachtet wird (vgl. Jakobs 2006, S. 319f).

Auch wenn Jakobs ihre Forschung primär auf textproduktive Anforderungen und Handlungskontexte richtet, gibt es vermutlich keinen Grund nicht auch ähnliches für die textrezeptive Seite anzunehmen, da der Prozess des Schreibens, wie sie selbst ebenfalls ausführt auch oft textrezeptive Teile beinhaltet.

Auch Weissenberg (2010) hält als Erkenntnis einer Analyse verschiedener Berufe die folgenden sprachlichen Anforderungen fest (eindeutige textgebundene Leseanlässe an dieser Stelle blau unterstrichen):

- Mündliche/schriftliche Arbeitsaufträge/Arbeitsanweisungen erkennen/verstehen
- Erläuterungen verstehen
- Mitteilung (über den Lagerbestand) machen
- Entnahme von Materialien dokumentieren
- Neues Material bestellen
- Material kontrollieren und reklamieren
- Bedienungsanleitungen verstehen
- Eine Mitteilung über einen Produktionsfehler machen

¹⁷ n pro Berufsgruppe = 6; 25 Berufsgruppen: Bankkaufleute, Beamte, Journalisten, Krankenschwestern, Kriminalbeamte, Mediziner, Notare, Öffentlichkeitsarbeiter, Rüstungsoffiziere, Steuerfachbearbeiter, Technische Redakteure, Wissenschaftler, Betriebs-/Sonderpädagogen, Ingenieure, Psychologen, Feuerwehrleute, Rettungssanitäter, Ergotherapeuten, Lehrer, Pfarrer, Polizisten, Kindergärtner, Krankengymnasten, Architekten und Postmitarbeiter.

- schriftliche Dokumentationen der Qualitätskontrolle/-sicherung schreiben und lesen

Auch wenn die übrigen Anforderungen nicht eindeutig auf Leseprozesse eukurrieren, so lassen sich doch eindeutig textbasierte kommunikative Handlungen erkennen.

Efing (2013) fasst in seiner Forschung zusammen, dass sofern unter einem Text ein lineares, kohärentes, komplex versprachlichtes Gebilde („Fließtext“) verstanden wird, Auszubildende scheinbar nur selten im Betrieb Texte produzieren oder rezipieren. Wenn jedoch unter den Begriff „Text“ auch depiktionale und hybride Repräsentationsformate subsumiert werden, können nicht wenige ausbildungsrelevante, typische, schriftliche Aufgaben, Textsorten und Darstellungsformen im Bereich „Lesen“ gefunden werden.

Dabei nennt er folgende(vgl. Efing 2013, S. 130):

- Kurznotizen
- Arbeitsaufträge/-anweisungen
- Artikel aus Fachzeitschriften
- Schulungsunterlagen
- Technische Zeichnungen, Schaltpläne, Datenblätter
- (Bedienungs-, Programmier-, Montage-, ...) Anleitungen
- An- und Unterweisungen (Sicherheitshinweise, Anleitungen zum Führen des Berichtsheftes, Projektleitfäden, ...)
- Bestellkataloge
- Fach-/Handbücher (Tabellenbücher, ...)
- Unternehmensbroschüren (Eigendarstellungen, Leitfäden für die Ausbildung)
- Juristische Dokumente (Arbeitsordnungen,...)

Efing hält daneben weiterhin fest, dass meist mehrere Textsorten vernetzt auftreten und dass die sprachwissenschaftliche Analyse dieser Textsorten, die sich im Bereich Lesen als relevant herausstellen, zeigt, dass es sich vorwiegend um Texte handelt, die in ihrer Formatierung zu großen Teilen stark normiert und standardisiert sind, was sich oft insbesondere in der Darstellungsform als Formular oder Tabelle (z.B. Berichtsheft, Messprotokoll, Arbeitsplan, Arbeitsauftrag, Liste) niederschlägt. Ferner sind diese Quellen durch eine ausgeprägte Multikodalität, d.h. Kombination aus Text, Bild, Grafik, Tabelle, Diagramm... – oder wie in dieser Arbeit bezeichnet: hybride Repräsentationsformate –, und rudimentäre Versprachlichung (Stichworte) gekennzeichnet. Der Umgang mit diesen Informationsquellen beschreibt er als Informationsentnahme und Beziehungsherstellung aus verschiedenen Darstellungsmodi und -kodes (welche vergleichbar mit Repräsentationsformaten sind) und Transformation in Handlungen (vgl. reading-to-

do), was – bezogen auf die Anforderungen – eine hohe kommunikative Kompetenz erfordert (vgl. Efing 2012, S. 9, 2013, S. 130f).

Neben der betrieblichen Realität stellt sich auch die Frage nach der schulischen Realität. Der schulische Teil der Ausbildung stellt meist noch zusätzliche/andere Anforderungen an die Auszubildenden. Das häufigste Medium mit dem Schüler konfrontiert werden ist das Schulbuch und der kompetente Umgang mit diesem, d.h. das Verstehen der jeweiligen Texte ist eine wichtige Voraussetzung für das erfolgreiche Absolvieren einer Ausbildung (vgl. Wengel 2013, S. 156).

Doch auch bei der Betrachtung von schulischen Anforderungen wird seitens der Forscher die Erfahrung gemacht, dass Fachtexte unterschiedlicher Fächer durch die besonderen Eigenschaften der Fachsprachlichkeit unterschiedliche Anforderungen an die Lesekompetenz stellen (vgl. Wengel 2013, S. 159), was sich auch bei der Betrachtung verschiedener Curricula (als nicht repräsentative Beispiele wurden Asphaltbauer, Gerber, Augenoptiker und Bestattungsfachkraft herangezogen) nochmals bestätigt (vgl. Wengel 2013, S. 161).

In diesem Kontext besonders erwähnenswert ist die Feststellung, dass bei Ausbildungsgängen mit geringeren textrezeptiven Anforderungen eine deutliche Spezifizierung auf bestimmte Repräsentationsmodi stattfindet: „Die geforderten Fähigkeiten beziehen sich also ausschließlich auf nicht-kontinuierliche Textsorten.“ (vgl. Wengel 2013, S. 162). Insgesamt kann auch hier als Fazit festgehalten werden, dass erste annähernde Betrachtungen zeigen, dass insgesamt in den letzten Jahren die Bedeutung kommunikativer Fähigkeiten zuzunehmen scheint (vgl. Wengel 2013, S. 166f).

All dies macht deutlich, dass arbeitsbezogene Lesefähigkeit einer der wichtigsten Typen funktionaler Lesefähigkeit ist, dabei aber bisher nur unzureichend erforscht wurde.

Es bleibt ferner die Frage, ob sich diese Beschreibung und Analyse der Leseanforderungen als problematisch für die betreffenden Individuen erweist – dem widmet sich das nächste Kapitel.

2.1.3. Vergleich von Anspruch und Leistung beruflicher Lesekompetenz

Nachdem gezeigt wurde, dass Lesen auf verschiedensten Ebenen unmittelbar mit beruflichen Handeln verbunden ist, stellt sich die Frage, ob die individuellen Akteure die an sie gestellten Anforderungen erfüllen können. Dabei erfolgt im Nachkommenden keine systematische Analyse des Erkenntnisstandes, sondern vielmehr ein exemplarisches Eingehen auf unterschiedliche Studien und Analysen, in denen Defizite bezüglich der Lesekompetenz von Schülern im beruflichen Bildungssystem ermittelt bzw. zusammenfassend dargestellt werden, da die Defizite in Studien wie PISA und vor allem ULME hinlänglich dokumentiert sind.

Für Berufsschüler zeigen Rexing et al. (2013) durch Analyse vorliegender Studien auf, dass der Hauptproblempunkt ist, dass es den allgemeinbildenden Schulen nicht zu gelingen scheint, auf

verstehendes Lesen im Kontext funktionaler Lesekompetenz adäquat vorzubereiten, wobei die Haupt- und Realschüler als besondere Risikogruppe identifiziert werden können (vgl. Rexing et al. 2013, S. 51).

Die zentrale Erkenntnis der ULME Studie ist nach Rexing et al. (2013, S. 49), dass die meisten Berufsschüler grundlegende Informationen aus einem Text entnehmen können, jedoch für viele Schüler große Schwierigkeiten bei anspruchsvolleren Verständnis- und Dekodierungsleistungen bestehen (vgl. Lehmann et al. 2006, S. 50).

Die Ergebnisse der im Rahmen der Implementation ausgewählter Materialien des Kölner Modellversuchs Leseförderung in der Berufsbildung gestellten Diagnostik der Lesekompetenz in den Eingangsklassen des Bau- und Ausbaugewerbes mittels des „Lesetests für Berufsschüler/-innen LTB“ (Drommler et al. 2006) zeichnen ein weiteres beunruhigendes Bild: 61 % der Berufsschüler erreichen nicht Kompetenzstufe I und sind folglich nicht in der Lage, explizite Einzelinformationen eines Textes, eines Diagramms oder einer Tabelle zu erfassen, verbunden mit den Berufsschülern, welche lediglich Kompetenzstufe I erreichen (32%), verfügen somit 93% der getesteten Berufsschülerinnen über eine Lesekompetenz, die im Kontext beruflicher Anforderungssituationen vermutlich nicht ausreichend oder zumindest als problematisch einzuschätzen ist (vgl. Rexing et al. 2013, S. 50f)¹⁸.

Zum Abschluss liefert auch die Analyse des Modellversuchs VERLAS (Verknüpfung von berufsfachlichem Lernen mit dem Erwerb von Sprachkompetenz und mathematisch-naturwissenschaftlicher Grundbildung) bei Berufsschülern in den Eingangsklassen dualer Bildungsgänge kein anderes Bild und weist auf ein scheinbar grundsätzliches Problem bezüglich berufliche Handlungsfähigkeit hin, denn für alle Berufsschüler wurden Schwierigkeiten im Bereich des Leseverständnisses von Sachtexten, Statistiken und Tabellen festgestellt (Rexing et al. 2013, S. 51). Auch wird darauf verwiesen, dass Ausbildungsbetriebe vor allem auf Defizite beim Lesen von Betriebsanweisungen und Vorschriften aufmerksam machen (Rexing et al. 2013, S. 51).

Die Konsequenzen werden zusammengefasst:

„Für das berufliche Bildungssystem bedeutet dies, dass an der ersten Schwelle viele Schülerinnen nur sehr einfache Leseaufgaben in angeleiteten Kontexten bewältigen können, was für eine Vielzahl von Leseaufgaben in Ausbildungs- und Alltagssituationen nicht ausreichend ist“ (Rexing et al. 2013, S. 48).

Auch Efing hält als Fazit fest, dass die von ihm untersuchten Berufsschüler¹⁹ bezüglich der Lesekompetenz Probleme mit dem Aufbau eines mentalen Modells eines Textes, mit der Verknüpfung verstreuter Informationen, dem Erkennen von impliziten Informationen an nicht exponierten Stellen und der Trennung zwischen Textinformation und Vorwissen haben. Daneben

¹⁸ Wobei (Rexing et al. 2013, S. 50f) auch explizit auf anzunehmende testtheoretische Konstruktionsmissstände des LTB hinweisen, die das Ergebnis beeinflusst haben könnten.

¹⁹ Dargestellt werden die mittels eines auf 90 Minuten angelegten Problemtypentests erhobenen Daten; eingesetzt bei 415 Schülern in über 10 Berufsfeldern v.a. aus der Grundstufe der berufsbildenden Schulen, 55% waren männlich; 31.8% hatten einen nicht-deutschsprachigen Hintergrund.

war eine weitere Problematik die Motivation und Leistungsbereitschaft seitens der Jugendlichen (vgl. Efin 2006, S. 44).

Dafür nennen Rexing et al. als mögliche Erklärungsansätze für diese und ähnliche Befunde (Rexing et al. 2013, S. 52):

geringes fachliches Vorwissen, wodurch das Zusammenspiel zwischen sprachlichem und fachlichem Wissen eingeschränkt wird,

ein geringer Wortschatz insbesondere im berufsbildungssprachlichen Kontext mit besonderen Problemen bei Lernenden mit Migrationshintergrund auf Grund eines häufig nur unsystematischen Erwerbs des Deutschen als Zweitsprache,

Probleme beim Zugriff auf im Gedächtnis gespeicherte Wörter bzw. Wortbestandteile,

defizitäres kognitives und metakognitives Strategiewissen, d.h. insbesondere fehlende Strategien beim Lesen und Behalten von Texten,

eine teilweise fehlende oder geringe Motivation zum Lesen bzw. ein negatives Selbstkonzept im Hinblick auf die eigene Lesekompetenz,

sowie mit besonderem Gewicht: die für Berufsschüler offensichtlich nicht erkennbare Relevanz von Lesekompetenz für die eigene berufliche Entwicklung

Biedenbach (2006) argumentiert dahingehend, dass im Jahr 2004 bundesweit 21% der Auszubildenden ihren Ausbildungsvertrag vorzeitig gelöst haben, davon laut ihrer Aussage viele sogar noch im dritten Lehrjahr und von denjenigen, die ihre Ausbildung abschließen wollten, haben 14,3% die Abschlussprüfung nicht bestanden. Daraus schließt sie, dass sich viele Prüflinge mit dem Prüfungsstoff überfordert fühlen und/oder den an sie gestellten Anforderungen nicht gewachsen sind. Sie folgert weiterhin, dass es daher vielen Jugendlichen an den grundlegenden Fertigkeiten für den erfolgreichen Berufsabschluss – insbesondere der Sprachkompetenz – mangelt (vgl. Biedenbach 2006, S. 17). Auch wenn die von Biedenbach gezogenen Schlussfolgerungen etwas übergeneralisierend und dramatisierend scheinen, lassen sich doch bedenkliche Tendenzen erkennen.

Auch Jordan (2011) nimmt an, dass das beschriebene Scheitern der Auszubildenden innerhalb der Berufsausbildung darauf zurückzuführen sei, dass immer mehr Jugendliche, wenn sie ihre berufliche Erstausbildung beginnen, nicht über die notwendigen sprachlich-rezeptiven Fähigkeiten verfügen, um sich das notwendige berufliche Fachwissen überhaupt aneignen zu können (vgl. Jordan 2011, S. 7).

Die von Rexing et al. (2013) geäußerte Kritik an den eingesetzten Tests offenbart jedoch auch wiederum die bisher problematische diagnostische Lage im Bereich der Lesekompetenzmessung in beruflichen Handlungszusammenhängen, da es kaum geeignete Lesekompetenztestes gibt, die sowohl die strukturellen Dimensionen der Repräsentationsformate und deren Einfluss auf kognitive Verarbeitungsprozesse berücksichtigen als auch die besondere Verwendungssituation

des zu lesenden Materials als Integrationsleistung in eine konkrete berufliche Handlung abbilden.

2.1.4. Leseförderung in der beruflichen Bildung

Die gesamte Tragweite des Beschriebenen kristallisiert sich in dem Punkt Leseförderung in der beruflichen Bildung. All das vorher genannte – insbesondere die Differenz zwischen notwendiger und tatsächlicher Leistung beruflicher Schüler bezüglich des Lesens – kann zwar als prekär eingestuft werden, dennoch ist die eigentliche Problematik die fehlenden Erfolge einer sonst scheinbar elaborierten Leseförderungsstrategie.

Dazu halten Ziegler/Gschwendtner (2010) fest, dass durch Interventionsforschung zu „Reciprocal Teaching“ an beruflichen Schulen bei der Leseförderung keine Effekte auf Klassenebene festgestellt werden konnten (vgl. Ziegler/Gschwendtner 2010), obwohl die metaanalytisch berichteten und global positiven Befunde, ebenso wie die Rezeption und unterrichtsbezogenen Empfehlungen des Ansatzes in der scientific community, suggerieren, dass dieser Ansatz zur Förderung des Leseverständnisses geeignet sei (vgl. Gschwendtner 2012, S. 246).

„Reciprocal Teaching [kann dabei] als ein Instruktionsansatz verstanden werden, dessen direkte Förderziele die Entwicklung lesebezogenen metakognitiven Wissens, metakognitiver Prozesse und metakognitiven Bewusstseins und darüber hinaus die Entwicklung des Leseverständnisses sind.“ (Gschwendtner 2012, S. 246)

Weiterhin führt Gschwendtner (2012) auch aus, dass die Befundlage zu Reciprocal Teaching insgesamt inhomogen ist. Die Dokumentation und Kausalität der Wirksamkeit scheint an viele noch unbekannte Variablen und Konstellationen geknüpft wie bspw. die Definition von Variablen und deren Operationalisierungen sowie eine Übertragung in ökologisch valide Schul- bzw. Klassenkontexte. Insbesondere hebt er hervor, dass die eingeforderten Lehrfunktionen im Zusammenhang mit regulären Schüler-Lehrer-Relationen, der heterogenen Schülerschaft mit zum Teil ungünstigen Bildungsbiographien starke Implementations- und Wirksamkeitsbarrieren darstellen können (vgl. Gschwendtner 2012, S. 246).

Bedeutsam ist in diesem Kontext vor allem, dass die bei leseschwachen Schülern erzielten Effekte nur wenig nachhaltig sind. Es konnte beobachtet werden, dass je länger die Trainingszeit war, desto geringere Effekte zu verzeichnen waren und das Training motivational nur als wenig wirksam eingestuft werden konnte. Als Kernursache dafür wird vermutet, dass es an fehlender wahrgenommener Bedeutung des Trainings für beruflich-betriebliche Kontexte liege (vgl. Ziegler/Gschwendtner 2010).

Daraus folgt, dass Förderprogramme prinzipiell an exakt diesem beruflich-betrieblichen Kontext ansetzen sollten, um die Chancen für eine erfolgreiche Intervention und nachhaltige Verbesserung zu erhöhen. Diesbezüglich konnte bereits aufgezeigt werden, dass allgemeines Lesetraining nicht so große Auswirkungen auf das arbeitsbezogene Lesen hat, wie arbeitsbezogenes Lesetraining (vgl. Mikulecky 1988).

Keimes/Rexing (2011a) halten in ihrer Analyse beruflicher Leseförderkonzepte alsdann fest, dass bis auf die genannten Interventionsstudien die Befunde sehr heterogen und vornehmlich empirisch wenig fundiert sind. Auffällig ist nach den Analysen von Keimes/Rexing, dass es bezüglich der Implementation von Leseförderung in die berufliche Bildung vornehmlich Modellversuche gibt, deren bemerkenswertes Charakteristikum ist, dass diese den Schwerpunkt wesentlich auf die *Konzeptentwicklung* legen, weniger auf die *Überprüfung der Wirksamkeit und Nachhaltigkeit der Konzepte*. Dabei nennen sie bspw. Modellversuche, welche mehrfach das Problem der Leistungsheterogenität von Berufsschülern anführen und daraus die Notwendigkeit der Differenzierung bzw. der individuellen Förderung ableiten. Wie diese jedoch konkret-praktische umgesetzt und organisiert werden kann und wie sich alsdann Maßnahmen in diesem Kontext real auf die individuelle Entwicklung der Lesekompetenz auswirken (können), wird nicht präzisiert (vgl. Keimes/Rexing 2011a, S. 91). Somit folgern sie: „Vor dem Hintergrund dieser eher fragmentarischen Befundlage kann die Notwendigkeit einer systematischen empirischen Begleitung von Förderansätzen abgeleitet werden.“ (Keimes/Rexing 2011a, S. 91)

Das Dargestellte besitzt jedoch weiterhin nicht nur bezogen auf Jugendliche Relevanz, sofern bedacht wird, dass Lesekompetenz eine trainierbare Fähigkeit ist, welche bei mangelndem Einsatz auch wieder degenerieren kann. So wurde festgestellt, dass größere Kompetenzverluste insbesondere bei Langzeiterwerbslosen, d.h. Erwerbslose, die 12 Monate und länger keiner Erwerbstätigkeit nachgegangen sind, in Bezug auf das Lesen zu verzeichnen sind (vgl. Rammstedt 2013, S. 17).

Somit ist die Notwendigkeit weiterer Forschung bezüglich der Leseanforderungen, -diagnostik und -förderung in beruflichen Handlungskontexten evident.

2.1.5. Rekapitulation des Untersuchten I

Zusammengefasst lassen sich acht zentrale Aussagen festhalten:

1. Lesen ist, ob aus individueller, gesellschaftlicher oder ökonomischer Perspektive besehen, außerordentlich wichtig.
2. Lesen am Arbeitsplatz erfordert zum Großteil insbesondere die Verwendung multipler Dokumente der drei Repräsentationsformate depiktional, deskriptional und hybride Variationen.
3. Lesen in beruflichen Kontexten scheint hauptsächlich in konkrete Arbeits- und Handlungszusammenhänge integriert zu sein (reading-to-do-literacy/funktionale Lesekompetenz).
4. Die Leseanforderungen am Arbeitsplatz sind vermutlich berufsübergreifend nicht unerheblich und werden wahrscheinlich noch steigen.

5. Die Leseanforderungen sind innerhalb eines Berufes und über verschiedene Berufsfelder hinweg sehr heterogen.
6. Deutsche Schüler scheinen den gestellten Anforderungen nicht angemessen gewachsen zu sein.
7. Etablierte Leseförderprogramme/-strategien zeigen nicht die erwarteten bzw. erwünschten Effekte bei schwächeren Schülern des berufsbildenden Bereichs.
8. Es fehlen bezüglich beruflicher Lesediagnostik und -förderansätze systematische empirische Untersuchungen und daraus abgeleitete Reflexionen.

3. Theoretische Ausarbeitungen

3.1. Lesen – Begriff, Situation, Material, Modellierung und Messung

3.1.1. Annäherung an das Konstrukt „Lesekompetenz“

In dieser Arbeit werden verschiedentlich die Begriffe (Lese-)Kompetenz, (Lese-) Fähigkeit, (Lese-)Fertigkeit verwendet, um die (erfolgreichen) Prozesse des Decodierens und Interpretierens von Informationen aus depiktionalen, deskriptionalen und hybriden Repräsentationsformaten zu beschreiben. Der Begriff Lesekompetenz wird seit Publikation der Ergebnisse aus den PISA-Untersuchungen quasi inflationär verwendet, ohne dabei in den jeweiligen Veröffentlichungen konkret definiert zu werden. Genau betrachtet besteht der Begriff aus drei Facetten. Einerseits aus dem Konstrukt „Lesen“ andererseits aus dem Begriff „Kompetenz“ und deren Kompositum „Lesekompetenz“.

Der Kompetenzbegriff im Allgemeinen

Zunächst wird der für das Verständnis zentrale Teilbegriff „Kompetenz“ expliziert, der seit seiner Einführung in die Erziehungswissenschaft in den 70er Jahren durch Heinrich Roth (vgl. Roth 1971, S. 180) umstrittener Gegenstand von Diskussionen ist²⁰, um ihn später über den Kontext des Lesen zu differenzieren (siehe unten).

Kompetenz ist in der deutschen Alltagssprache nicht erst seit den auch in Deutschland groß angelegten Erhebungen von Schülerleistungen (large scale assessments) ein vertrautes Wort. Hartig (2008, S. 15) führt an, dass laut einer Zählung des Projekts Deutscher Wortschatz im Jahr 2006 „Kompetenz“ zu den 5.000 am häufigsten verwendeten Wörtern der deutschen Sprache gehört. Die Bedeutungen und Verwendungen des Begriffs in der Alltagssprache sind derart vielgestaltig, dass der Begriff in einigen Kontexten zusätzlich synonym zu „Zuständigkeit“ im

²⁰ Vgl. ausführlich und kontrovers: Dilger (2004), Klieme (2004), Gnahn (2007), Körber (2007), Klieme und Hartig (2007), Lauth (2010), Mugerauer (2012), Grabowski (2014)

juristischen und/oder organisatorischen Sinne verwendet wird. Auch wird der Begriff häufig mit „Fähigkeit“ assoziiert. Diese dargelegte Alltagsverwendung und, dass in den mit Bildungsprozessen und deren Ergebnissen befassten Fächern und Teildisziplinen keine konsequente und einheitliche Festlegung des Begriffs existiert, macht an dieser Stelle eine inhaltliche Definition des notwendig (vgl. Hartig 2008, S. 15f).

Zunächst ist festzuhalten, dass „Kompetenz“ ein sozial konstruierter Begriff zur Beschreibung von etwas ist (vgl. Vonken 2011, S. 22). So lässt sich die Vielfalt der verschiedenen widersprüchlichen Kompetenzbegriffe erklären, denn Vonken (2011) führt aus: „dass hinter Kompetenz nicht eine Armada abendländischer Philosophie und Religion steht, die wie bei ‚Bildung‘ schlagkräftig die Richtung angibt.“²¹ (Vonken 2011, S. 23) So dass der Begriff flexibel besetzt werden kann.

Bohlinger (2007/2008, S. 118) differenziert die Hauptverständnisse des Begriffs Kompetenz grundsätzlich in fähigkeits- oder tätigkeitsbezogene Ansätze, d.h. einerseits Ansätze, die Kompetenz als Fähigkeit zur Bewältigung von Situationen verstehen. Diese Richtung führt sie zurück auf die psychologische Theoriebildung und verortet deren Anwendung vor allem im Kontext der Entwicklung von Handlungskompetenz. Teilweise uneindeutig ist bei diesen Strömungen, ob es sich bei Kompetenzen um ein Persönlichkeitsmerkmal oder um eine durch Bildungs- und Erziehungsprozesse erworbene Fähigkeit handelt. Andererseits existieren Ansätze, die die Erzeugung und Ermöglichung von Situationen fokussieren. Diese stehen in einer gesellschaftskritischen Theorietradition, in denen Kompetenz als Medium für gesellschaftliche Transformationsprozesse betrachtet wird.

Groeben (2002) definiert Kompetenzen als das individuelle Potential dessen, was eine Person unter optimalen Umständen leisten könnte. Andererseits zeige sich dieses Potential in konkreten Situationen als spezifisches Verhalten bzw. Handeln. Sodann umfasse das Kompetenzkonzept sowohl die Ebene der Fertigkeiten als auch die der Fähigkeiten. Fertigkeiten seien dabei konkrete, situations- und aufgabenbezogene Verhaltensweisen, Fähigkeiten hingegen universelle, generelle sowie stabile Konstrukte (vgl. Groeben 2002, S. 13).

Problematisch an der vielseitigen in verschiedenen wissenschaftlichen Kontexten verbreitete Verwendung ist, dass derartig uneinheitliche und sich teilweise gegenseitig widersprechende oder ausschließende Verständnisse von wissenschaftlichen Konstrukten auf der einen Seite Kommunikationsschwierigkeiten evozieren können, aber auch auf der anderen Seite den Vergleich und die Bewertung unterschiedlicher empirischer Befunde (je nach dem welchem Verständnis von Kompetenz gefolgt wird) quasi wirkungslos werden lassen (vgl. Hartig 2008, S. 16).

Gelegentlich wird daher konstatiert, dass Kompetenz an sich nicht definierbar sei oder gar keiner Definition bedarf (vgl. Wollert 1997, S. 328 zitiert nach Vonken 2011, S. 27).

²¹ Ob jedoch der Bildungsbegriff (aus diesem Grund bzw. überhaupt) einheitlicher und präziser definiert ist, sei dahingestellt.

Historisch betrachtet findet sich der Ursprung des Begriffs im Lateinischen – einerseits als „competens“ in der Bedeutung „angemessen“ bzw. andererseits als „competere“ was mit „zusammentreffen“ übersetzt werden kann. Im ersteren Sinne wurde der Begriff ins Deutsche übertragen und im 19. Jahrhundert bspw. in der Rechtsphilosophie als Synonym für „Zuständigkeit“ verwendet oder von Max Weber in seiner Untersuchung von Typen der Herrschaft. Wollersheim expliziert die zweite Bedeutung, in dem er ausführt: „wenn die Erfordernisse der Situation mit dem individuellen Konglomerat von Fähigkeiten und Fertigkeiten eines Menschen zusammentreffen, so besitzt dieser also die Kompetenz zur Bewältigung der Situation“ (Wollersheim 1993, S. 89 zitiert nach Vonken 2011, S. 21).

Auch wenn der Ursprung der Verwendung in den Erziehungswissenschaften nicht letztgültig geklärt ist, lässt sich die allgemeine Diskussion zurückführen auf Noam Chomsky und seine Theorie der Sprachkompetenz und Performanz als Replik auf Skinners behavioristische Sprachtheorie (Aebli 1980). Der Einsatz in der Erziehungswissenschaft folgt tendenziell der Bedeutung von Befähigung und Können (vgl. Vonken 2005, S. 18). Wobei schon früh eine Abgrenzung zum Konstrukt der Qualifikation angestrebt wurde. Der deutsche Bildungsrat definierte Kompetenz als „Lernerfolg im Hinblick auf den Lernenden selbst und seine Befähigung zu selbstverantwortlichem Handeln im privaten, beruflichen und gesellschaftlich-politischen Bereich“ (Deutscher Bildungsrat 1974).

Immer wieder hervorgehoben wird der in verschiedenen Kompetenzdefinitionen angelegte Bewertungsaspekt – wie bspw. in der Definition von Sloane/Dilger: „Kompetenzen sind die für die Bewältigung von spezifischen Aufgaben notwendigen Fähigkeiten und Fertigkeiten des Subjekts“ (Sloane/Dilger 2005, S. 6) bzw. „individuelle Fähigkeiten, die konkretes Handeln ermöglichen und empirisch aus der Beobachtung der jeweiligen Handlung und an der damit verbundenen Leistung gemessen werden können“ (Sloane/Dilger 2005, S. 7). Auch Freys Definition kann dazu gerechnet werden:

„Kompetenz [kann man] als ein Bündel von körperlichen und geistigen Fähigkeiten bezeichnen, die jemand benötigt, um anstehende Aufgaben oder Probleme zielorientiert und verantwortungsvoll zu lösen, die Lösungen zu bewerten und das eigene Repertoire an Handlungsmustern weiterzuentwickeln.“ (Frey 2004, S. 904)

Im Antrag für das DFG-Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ wird Kompetenz definiert als „kontextspezifische kognitive Leistungsdisposition, die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“ (Klieme/Leutner 2005, S. 4) und darüber hinaus „erlernbare und vermittelbare Leistungsdispositionen“ (Klieme/Leutner 2005, S. 4) sind. Später wird noch der Passus „unter Ausschluss motivationaler und affektiver Faktoren“ (Hartig/Klieme 2006, S. 129) ergänzt.

Dieser Definition wird im weiteren Verlauf der vorliegenden Arbeit gefolgt, daher wird sie noch weiter betrachtet. Es lässt sich fragen, ob das Konzept „Kompetenz“ nicht bereits hinreichend

mit etablierten Begriffen beschrieben werden kann, insbesondere wenn ein so unscharfer Begriff wie der der Kompetenz verwendet werden soll. Nach Hartig (2008) ist gerade die Kombination des Kompetenzbegriffs mit spezifischen Kontexten eine eigständige Dimension. Genauso trägt das Kriterium der Erlernbarkeit zur Abgrenzung zu anderen Konstrukten, wie der Intelligenz²² bei. Auch die Exklusion von motivationalen oder affektiven Faktoren ermöglicht eine Unterscheidung zwischen kognitiven Leistungsvoraussetzungen und motivationalem Einwirken (vgl. Hartig 2008, S. 18–20).

Beachtet werden muss jedoch, dass der für die Definition des Kompetenzkonstrukts relevante Kontext ebenfalls problematisch sein kann, sofern dieser zu eingeschränkt angelegt ist, denn dann ließe sich für jeden beliebigen Testinhalt eine Kompetenzdefinition konstruieren, in der die relevante Situation schlicht mit bestimmten Aufgabeninhalten gleich gesetzt wird. Daher muss der relevante Kontext für die Definition hinreichend konkret sein, andererseits sollte er auch nicht zu eng gefasst sein, da andernfalls simples Sachwissen oder isolierte Fertigkeiten zu Kompetenzen avancieren. Ein dafür geeignet scheinendes Kriterium ist der Bezug auf eine Menge realer Situationen, in denen bestimmte, analoge Anforderungen bewältigt werden müssen – Hartig führt dazu aus, dass real dabei pragmatisch als außerhalb des Bildungsprozesses verstanden werden kann (vgl. Hartig 2008, S. 20f).

So dann ermöglicht ein derartiges Verständnis von Kompetenz „eine Unterscheidung von Wissen und dessen Anwendung“ (Hartig 2008, S. 21).

Dieser Kontext wird hier konkret spezifiziert durch den anderen Bestandteil des Kompositums – der Begriff des Lesens. Dieser wird daher im Zusammenhang mit dem Konstrukt Lesekompetenz im nächsten Kapitel näher beleuchtet.

Lesen und Lesekompetenz

Der Begriff „Lesen“ ist nicht weniger uneindeutig als der vorhergehende Begriff der Kompetenz.

Stein (2010) hält diesbezüglich fest, dass es keinen Mangel an Darstellungen zur Geschichte der Schrift, der Schreibwerkzeuge und Schriftträger, der Bibliotheken und des Buchhandels, der Literatur und der Autoren existiert, die Situation bezüglich Lesen und der Lektüre sich jedoch anders darstelle. Er führt aus, dass der Eindruck entstehen könne, als ob das Lesen eine zeitlose Fähigkeit sei, die lediglich in den verschiedenen Epochen bei unterschiedlichen Personen in unterschiedlicher Intensität vorkäme. Was unter anderem auch daher rühre, dass Lesen eine menschliche Tätigkeit sei, die der Schrift vorausginge. Sie sei sozusagen genetisch älter, denn schon lange bevor vor über 5000 Jahren die Schrift in Gebrauch genommen wurde, hatten sich im Zuge der phylogenetischen Entwicklung des Gehirns spezifische Regionen gebildet, welche für Sprache zuständig seien und auf deren Aktivität die Kompetenz für Lesen und Schreiben aufbaue. Des Weiteren gelte Lesen als die ältere Kompetenz, da sie sich neurobiologisch aus der

²² Im Sinne einer allgemeinen kognitiven Grundfähigkeit

Fähigkeit des Spurenlesens entwickelte, die schon den Homo habilis und Homo erectus in den Stand versetzten, sein Überleben zu sichern (vgl. Stein 2010, S. 253f).

Historisch betrachtet bedeutet „lesen“ bis ins Mittelhochdeutsche „auswählen, sammelnd aufheben“. Ergänzend wurde es zu dem Zeitpunkt auch für „erzählen, berichten“ und für das Zusammenstellen von Buchstaben verwendet und weiterhin entwickelte es sich zu einem Begriff für die Sinnentnahme aus geschriebenen und gedruckten Sprachzeichen (vgl. Fritz/Suess 1986, S. 12). Etymologisch lässt sich die Entwicklung des Begriffs durch germanische Kulturverhältnisse begründen: „Lesen“ war ursprünglich der Begriff für das Auflesen der zur Weissagung ausgestreuten Runenstäbchen und beschrieb gleichzeitig die Bedeutungsentnahme also das Enträtseln von Runen (vgl. Fritz/Suess 1986, S. 12).

Fritz/Suess (1986) beobachten, dass „Lesen“ kein autonomer Begriff bzw. keine autarke Tätigkeit ist, denn immer ist ein Objekt (Medium, Text) vorhanden, welches „gelesen wird“ und ein Subjekt (Leser), das „liest“ (vgl. Fritz/Suess 1986, S. 12f).

Gesellschaftliche und individuelle Priorität gewann das Lesen während der Entwicklung der Kulturgeschichte der Literalität auf Grund der Tatsache, dass es bis ins 19. Jahrhundert vor dem Schreiben gelernt wurde. Eine der Folgen war, dass viele Menschen besser lesen als schreiben bzw. nur lesen konnten. So dann schlussfolgert Stein (2010), dass die Teilkompetenz Lesen, vom mühsamen Buchstabieren können über das Mitlesen bis zum lauten bzw. stummen Lesen, eine Vielfalt von Partizipationsmöglichkeiten an der Schriftkultur ermögliche, ohne dass diese von der Fähigkeit, schreiben zu können, abhängen (vgl. Stein 2010, S. 254).

Posterior wird Lesen jedoch unter der Perspektive betrachtet, dass Lektüre der bloße Vollzug des Geschriebenen sei, d.h. eine Technik des Ablesens und Wiedergebens von Texten, was besonders dann geschieht, wenn das Geschriebene lediglich als Verschriftung von Gesprochenem begriffen wird, welches im lauten Lesen wieder zu Gehör gebracht wird. Dies ist allerdings ein sehr reduziertes und zudem hierarchisiertes Verständnis vom Verhältnis von Sprache und Schrift (vgl. Stein 2010, S. 254).

Jordan stellt fest, dass der Begriff des Lesen bis in die 80er Jahre des vergangenen Jahrhunderts teilweise synonym zum Begriff der Lesekompetenz verwendet wurde (vgl. Jordan 2011, S. 30).

Insgesamt „[ist] der Begriff Lesen nicht eindeutig. Normalerweise bezeichnet er eine individuelle Tätigkeit (oder eine allgemeine Fähigkeit), die das Rezipieren von geschriebener Sprache durch das Auge (Augenbewegung beim Lesen, Saccade) oder den Tastsinn (Blindenschrift) bezeichnet, wobei das Entnehmen von Bedeutung und Informationen im Mittelpunkt steht.“ (Glück 2000, S. 405)

Auch im fachwissenschaftlichen Gebrauch ist die Trennung von Chomskys dichotomischer Unterscheidung von Kompetenz – im Sinne einer allgemeinen Fähigkeit – und Performanz – im Gegensatz dazu als konkrete Tätigkeit – nicht immer eindeutig zu leisten. Weiterhin wird eine Spanne, d.h. eine Ungleichsetzung von Bedeutung hin zu Information eröffnet. Im Mittelpunkt stehen bedeutet freilich nicht, dass Lesen keine anderen Funktionen haben kann, sondern, dass

diese die hauptsächlich vorkommenden sind. Die Unterscheidung von Information und Bedeutung kann hier nicht en detail ausdifferenziert werden, es lässt sich aber festhalten, dass Informationsentnahme und Bedeutungsgenerierung nicht nur unterschiedliche Begriffe sind, sondern auch verschiedene Prozesse. Damit wäre der Bogen zum Begriff der Kompetenz geschlagen, indem ein sich in einer konkreten Situation zeigendes Verhalten zu unterscheiden ist vom bloßen Wissen über etwas.

Fritz und Suess verstehen „Lesen“ in Anlehnung an Aust (1983), als ein Werkzeug, um einer Botschaft den Sinn zu entnehmen (vgl. Fritz/Suess 1986, S. 13).

Eine der frühesten Definitionen des Kompositums „Lesekompetenz“ stammt aus einer Untersuchung bei 20jährigen Rekruten in der Schweiz. Dort wird Lesekompetenz definiert und gemessen als die Fähigkeit, verschiedenste gedruckte Kommunikationsformen zu bewältigen (vgl. Bonfadelli 1980, S. 191).

Ausgeprägt bei jedem Versuch einer Definition des Konstrukts Lesekompetenz ist das antinomische Spannungsfeld, in dem sich jede präzise und doch umfassende, jede operationalisierbare und doch allgemeingültige Begriffsbestimmung abstrakter kognitiver Prozesse bewegt. Auf der einen Seite kann die Definition von Fritz und Suess als sehr allgemein gehalten bezeichnet werden, denn einer Botschaft den Sinn zu entnehmen, kann sich auch auf mündliche Kommunikation oder komplexe soziale Situationen beziehen. Und wann der Sinn entnommen wurde, müsste weiterhin noch an Hand von beobachtbaren Kriterien spezifiziert werden.

In der pädagogischen Rekrutenprüfung der Schweiz hingegen wird stärker gedrucktes Material fokussiert und damit mediale Prozesse und Formate, die im heute (also knapp 40 Jahre später) genuin zum Lesematerial gehören (Hypertexte, Lesen am Computer,...) ausgeblendet. Interessant ist, dass hier eher der Produktionsprozess (drucken) als das Präsentationsformat betont wird. Dafür ist der Begriff Kommunikationsformen sehr weitgefasst und spätestens bei der Operationalisierung ist fraglich, ob ein Brief und eine Bedienungsanleitung die gleichen Anforderungen an den Rezipienten stellen, wie ein wissenschaftlicher Text oder ein Märchen. Ferner ist der Terminus bewältigen für eine Definition noch sehr schwammig, denn ein Text, in dem jedes Wort kurz betrachtet oder laut ausgesprochen wurde, ist einem bestimmten Verständnis folgend auch „bewältigt“ worden, vielleicht sogar verarbeitet worden – verstanden jedoch ist eine andere Dimension.

Auch innerhalb renommierter und hoch rezipierter Studien (IALS, IGLU, DESI und PISA) wird Lesekompetenz unterschiedlich definiert.

In der internationalen Erwachsenenstudie „International Adult Literacy Survey“ (IALS) (vgl. Notter et al. 1999) wird unter Lesekompetenz die Fähigkeit, gedrucktes Informationsmaterial aus dem beruflichen, privaten oder öffentlichen Alltag zu verstehen und zur Erreichung persönlicher Ziele oder zur Entwicklung weiterer Kompetenzen zu benutzen, verstanden (vgl. Stoll 1999, S. 26). Hierbei wird das Verständnis dahingehend erweitert, dass über eine Entzifferung

von Buchstaben hinausgegangen wird und sich vielmehr auf die Erschließung und Verwendung von Textzusammenhängen bezogen wird.

Auch hier zeigt sich wieder das angesprochene Spannungsfeld: Wenngleich der Kontext für Lesen bei IALS stark fokussiert wird, liegt doch der Schwerpunkt wieder auf gedrucktem Material, welches hier eingegrenzt wird als Informationsmaterial und in eine intentionale Struktur eingebettet wird. Ferner wird an dieser Stelle auf einen Verstehenskontext hingewiesen und so deutlich, dass eine Information aus einem Text zu entnehmen und ebendiesen Text zu verstehen zwei unterschiedliche Aspekte sind.

Das Konzept der Lesekompetenz in der IGLU-Studie von 2001 (international: PIRLS: Progress in International Reading Literacy Study) geht von einem interaktiven Modell des Leseverständnisses nach Rumelhart (1994) aus. Dabei wird Lesekompetenz als ein konstruktiver, interaktiver Prozess verstanden, in dem die drei Faktoren Leser, Text und Kontext interagieren (Bos et al. 2003, S. 71). Dementsprechend werden drei verschiedene Bereiche von Lesekompetenz unterschieden:

1. Der Aspekt der Verstehensleistung
2. Die Intention beim Lesen literarischer und expositorischer Texte
3. Sowie die Einstellungen und die Lesegewohnheiten des Lesers

Dabei ist die Verstehensleistung auf Seiten des Lesers abhängig von Dekodierfähigkeiten, allgemeinen kognitiven Voraussetzungen, motivationalen Aspekten, Wissen und Leseerfahrungen (vgl. Bos et al. 2003, S. 72).

Die DESI-Studie (2003/2004) definiert Lesen als eine spezifische, zielorientierte Form der rezeptiven Informationsverarbeitung, bei der sich die Leser sowohl auf die graphischen Zeichen und damit eine visuell orientierte Verarbeitungsweise stützt als auch auf sprachlich formale Kompetenzen und auf domänenspezifisches Weltwissen (vgl. Nold/Willenberg 2007, S. 23). Dabei gründet sich das verwendete Prozessmodell der Lesekompetenz auf zahlreiche Studien aus der kognitionspsychologischen Leseforschung. Demnach seien bei jedem Leseakt sechs differenzierbare Tätigkeiten zu unterscheiden, deren Komplexität bei zunehmender Textmenge ansteigt. Diese Teiltätigkeiten des Lesens sind: Informationen entnehmen, Inferenzen bilden, Wissen aktivieren, Verknüpfungen zwischen verschiedenen Absätzen und Textpassagen herstellen und abschließend mentale Modell Konstruktion (vgl. Garbe 2009, S. 28–30).

Die wohl verbreitetste Lesekompetenzdefinition ist die der PISA-Konzeption: Lesekompetenz ist die Fähigkeit geschriebene Texte unterschiedlicher Art in ihren Aussagen, ihren Absichten und ihrer formalen Struktur zu verstehen und sie in einen größeren sinnstiftenden Zusammenhang einzuordnen, sowie in der Lage zu sein, Texte für verschiedene Zwecke sachgerecht zu nutzen (vgl. Baumert et al. 2001b, S. 22).

Neben fortlaufend geschriebenen Texten sog. „kontinuierlichen Texte“ (deskriptionale Repräsentationsformate), werden auch „nicht-kontinuierliche Texte“, d.h. Darstellungen wie Diagramme, Bilder, Karten, Tabellen oder Grafiken (hybride oder depiktionale Repräsentationsformate) zunehmend einbezogen (vgl. Artelt et al. 2004, S. 141).

Auf Grund des vorher dargelegten lässt sich zumindest festhalten, dass Lesen kein passiver Prozess der Bedeutungsentnahme ist, sondern eine aktive Konstruktionsleistung des Individuums darstellt, bei der die verschiedenen im Text enthaltenen Inhalte aktiv mit dem Vor- und Weltwissen des Rezipienten in Verbindung gesetzt werden müssen. Weiterhin ist Lesen ein komplexer Vorgang aus mehreren flexiblen und kontextabhängigen Teilprozessen auf der Wort-, Satz- und Textebene (vgl. Kapitel 0, 0 sowie 0). Auf der untersten Ebene sind das Erkennen von Buchstaben und Wörtern und die Erfassung von Wortbedeutungen angesiedelt (Wortebene), auf der mittleren Ebene steht die Schaffung semantischer und syntaktischer Relationen zwischen Sätzen oder Phrasen im Vordergrund (Satzebene). Auf der oberen Ebene stehen dann die satzübergreifende Integration von Sätzen zu Bedeutungseinheiten und der Aufbau einer kohärenten mentalen Repräsentation der Bedeutung des Textes (propositionalen Textebene) im Fokus. Die kohärente mentale Repräsentation des gelesenen Materials oder des Sachverhalts (so genanntes Mentales Modell oder Situationsmodell) ist das Resultat einer Text-Leser-Interaktion, die mehr oder weniger intensiv sein kann, bzw. des Aufeinandertreffens von Textmerkmalen mit spezifischen Voraussetzungen des Lesers (bspw. Vorwissen, Intentionen, Lesetechniken). Die obengenannten relevanten Teilfähigkeiten und Prozesse des Lesens können teils als automatisiert angesehen werden, teils werden sie jedoch vom Leser auch wissend- und willentlich gesteuert (z.B. bei Verständnisproblemen, logischen Widersprüchen, spezifischen Zielen) (vgl. Artelt et al. 2007, S. 11f).

Es wird sodann für den weiteren Verlauf folgendes an PISA angelehntes Verständnis von Lesekompetenz zu Grunde gelegt: Lesekompetenz ist die Fähigkeit unterschiedliche literale Repräsentationen depiktionalen, deskriptionalen oder hybriden Formats bezüglich ihrer Aussagen, ihren Absichten und ihrer formalen Struktur zu verstehen und sie in einen konkreten Handlungszusammenhang sinnstiftend einzuordnen, sowie vor allem in der Lage zu sein, sie für diesen Handlungszusammenhang sachgerecht und erfolgreich zu nutzen. Der Grad der Kompetenz kann am Resultat und der Dauer des zu Erzielen gemessen werden.

Lesen ist der entsprechende mentale Prozess, unabhängig des als erfolgreich oder gescheitert zu bewertenden Ergebnisses.

Unter deskriptionalen Repräsentationen werden kontinuierliche Texte, die eine gewisse innere Kohärenz aufweisen, verstanden, unter depiktionalen Repräsentationen werden vornehmlich Graphiken, Diagramme und Pläne verstanden. In der Kategorie der Hybridformate werden diskontinuierliche Texte, wie Tabellen, Listen und Formulare zusammengefasst, sowie Repräsentationen, welche aus der Kombination zweier unterschiedlicher Formate bestehen und trotzdem eine Quelle bilden.

Aus Gründen der Lesbarkeit werden unter dem Begriff „Text“ diese unterschiedlichen Repräsentationsmodi (depiktional, deskriptional oder hybride Formate) zusammengefasst.

Dabei sollte ferner beachtet werden, dass in einer konkreten Lesesituation auch verschiedene Quellen benutzt werden können. Somit ist neben dem Format der einzelnen Quelle auch der gesamte Lesestimulus zu kategorisieren. Werden nur Texte eines Formats bspw. nur deskriptionale Repräsentationen verwendet, so ist der gesamte Stimulus als deskriptional zu verstehen, ebenso bei depiktionalen oder hybriden Formaten. Hervorzuheben ist diese Betrachtungsweise bei Quellen verschiedenen Formats: sobald zwei unterschiedliche Quellen verschiedener Modi, bspw. deskriptional und hybrid oder depiktional und deskriptional, in derselben Lesesituation verwendet werden, kann der Lesestimulus insgesamt als hybrid eingestuft werden.

Eine erste Annäherung an den Leseprozess und seine Modellierung, wie beides später ausführlich in Kapitel 3.3 dargelegt wird, kann durch die Betrachtung der verschiedenen auf den kognitiven Prozess einwirkenden Determinanten erfolgen.

Diese global zu berücksichtigenden Merkmale und Variable können im Rahmen eines Tetraedermodells beschrieben werden, in dem die relevanten Aspekte nach vier Merkmalsklassen gruppiert werden können (vgl. Jenkins 1979):

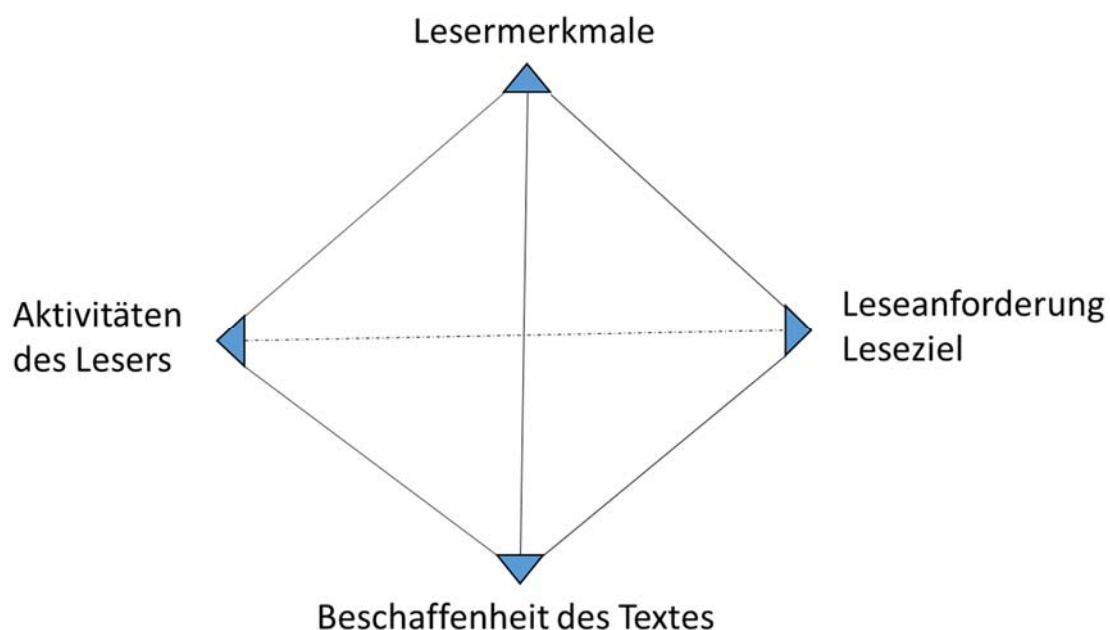


Abbildung 1 – Determinanten des Leseprozesses (Jenkins 1979)

Die Merkmale weisen in drei Punkten einen starken Leserbezug auf. Dies umfasst einerseits die konkreten Aktivitäten, die der Leser zur Texterschließung betreibt, wie bspw. der Einsatz von Textverstehensstrategien, andererseits die allgemeinen Lesermerkmale, wie kognitive Grundfähigkeiten, drittens die Leseanforderungen und -intentionen, die, wie aus Abbildung 1 zu erkennen ist, direkt auf die Leseraktivitäten wirken. Diese Leseintentionen und deren möglicher Ein-

fluss auf den Leseprozess werden eingehend in Kapitel 3.2.2 dargestellt und analysiert. Es werden in dem Modell aber auch textbezogene Merkmale wie Länge oder Schwierigkeit berücksichtigt, welche das im Einzelnen sein können wird in Kapitel 3.2.6 ausführlich beleuchtet.

In die Darstellung (Abbildung 1) könnte noch explizit der situative bzw. kontextuelle Rahmen hinzugefügt werden, so dass eine pentagonale Struktur entworfen werden kann, die gegenüber der ursprünglichen Konstruktion einen weiteren Pol hat (vgl. Abbildung 2).

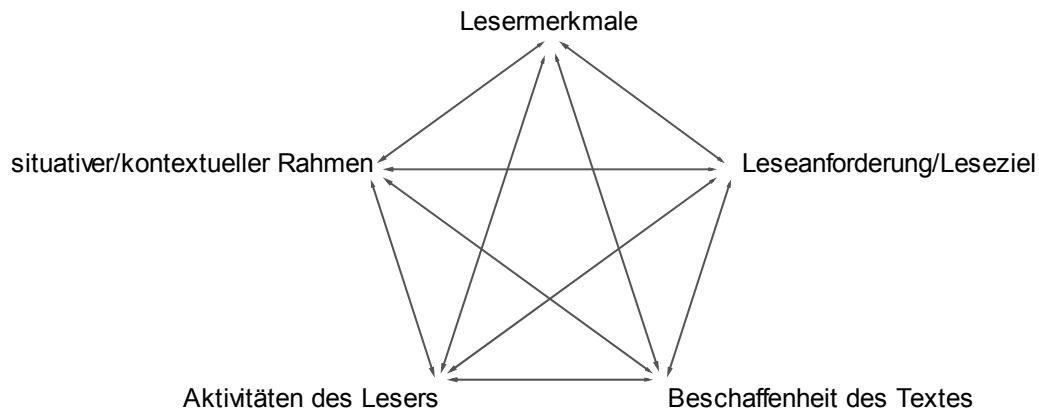


Abbildung 2 – Determinanten des Leseprozesses II

Dies scheint einerseits sinnvoll, da Ziele intrinsisch motiviert sein können, aber auch extrinsische Zielvorgaben existieren können. Diese können bspw. durch den kontextuellen Rahmen induziert werden. Auch kann das situative Setting bspw. dazu führen, dass nur eine begrenzte Zeit zur Verfügung steht, so dass der Leser andere Textbewältigungsstrategien aktiviert, als die, die er verwenden würde, wenn der Zeitfaktor keine Rolle spielt.

3.1.2. Kommunikationssituation und Leseintention

Bei der Betrachtung von Lesekompetenz sollten zwei Aspekte nicht außer Acht gelassen werden. Einerseits das besondere Verhältnis in die eine textbasierte Kommunikationssituation die beteiligten Parteien bringt, andererseits davon ausgehend die Basis dieses Verhältnisses – der Text – und die damit verknüpften Besonderheiten. Dies geschieht im Folgenden mit dem Fokus auf berufliche Kontexte.

Kommunikationssituation beim Lesen

Zunächst ist allgemein zu konstatieren, dass Textualität eine spezifische Form der nicht-natürlichen oder artifiziellen Sprachverwendung darstellt. Sie ermöglicht die Realisierung spezieller Funktionen des Kommunizierens, die mit der gesprochenen Sprache alleine nicht möglich sind. Es gibt Vermutungen, dass literale der Kommunikation der oralen bezüglich der kulturellen Basis überlegen sei, da orales Denken notwendig konkret und bildlich bleibt, im Gegensatz zum begriffliche gesehen logisch geordneten Denken einer literalen Kultur (vgl. Klicpera/Gasteiger-Klicpera 1995, S. 4; Havelock 1982 zitiert nach Fritz/Suess 1986, S. 96). Insbesondere der

Alphabetschrift wird eine derartige Förderung zugeschrieben, da diese leichter erlernbar ist und sowohl als effizienter als auch flexibler gesehen wird, da diese die kleinsten Lauteinheiten abbildet, statt bspw. größerer aber weniger definierter Silbenstrukturen (vgl. Fritz/Suess 1986, S. 96f). Jedoch konnten bisher keine eindeutigen Beweise dafür vorgelegt werden.

Anthropologen und Ethnologen konnten diesbezüglich bisher festhalten, dass komplexe Kommunikationssysteme auch im Tierreich vorkommen und somit das verbale bzw. gestische Kommunizieren nicht mehr als Primat des Menschen bezeichnet werden kann. Dem Gegenüber ist die geschriebene Sprache eine exklusiv menschliche Kommunikationsform. Diese ist historisch betrachtet im Vergleich zur gesprochenen Sprache eine relativ junge Erfindung, da systematische Schriftsysteme erst ca. 4.000-3.000 Jahre vor Christus entwickelt wurden (vgl. Scribner/Cole 1981 zitiert nach Fritz/Suess 1986, S. 95). Aus der konstruierten Neuheit der Schrift und gleichzeitig ihrer Universalität rührt auch der Disput über die Frage, ob Schreib- und Lesefähigkeit genetische Merkmale des Menschen sind oder ausschließlich antrainierte artifizielle Fähigkeiten. Problematisch an dieser Diskussion ist, dass Illiteralität häufig mit mangelnder Durchschnittsintelligenz oder fehlender Kultur gleichgesetzt wird (vgl. Havelock 1980, S. 41ff zitiert nach Fritz/Suess 1986, S. 95), was grade im Hinblick auf Buchkultur und Schriftlichkeit von einigen Forschern auch als eurozentrische Arroganz ausgelegt wird (vgl. Weidhaas 1984, S. 2051 zitiert nach Fritz/Suess 1986, S. 95).

Wenngleich sich derartig komplexe sozio-kulturelle Annahmen und Positionen nur schwerlich belegen lassen und in diesem Kontext zu weit führen, lässt sich nicht von der Hand weisen, dass sich die geschriebene und die mündliche Sprachverwendung bereits strukturell deutlich voneinander unterscheiden – Schrift ist nicht nur die simple Transformation des Gesprochenen in ein anderes Medium. Daher rührt auch die von Ethnologen häufig vorgenommene Trennung in primitive und zivilisierte Gesellschaften, an Hand der Einführung der Schrift in den jeweiligen Gesellschaften. Lévi-Strauss (zitiert nach Charbonnier 1973) bezeichnet diesen Schritt als die entscheidende Zäsur in der Entwicklung einer Gesellschaft, was auf Untersuchungen der Veränderungen, die durch diese Einführung bewirkt wurden, beruht.

Bereits die Kommunikationssituation ist grundverschieden: Während eines mündlichen Gesprächs besteht idealiter eine raumzeitliche Koinzidenz von Sprecher, Hörer und Signal. Dadurch verfügen beide Parteien nicht nur über einen gemeinsamen Referenzrahmen, in dem sie die sprachliche Situation approximativ gleich wahrnehmen, sondern auch über die Möglichkeit, auf ebendiesen mit verbalen und non-verbalen Mitteln direkt zu referieren, so dass die Möglichkeit besteht, etwaige Verständigungsprobleme unmittelbar zu signalisieren und ggf. auch auszuräumen. Mündliche Äußerungen weisen somit also stets eine empraktische Eingebundenheit auf (vgl. Günther 1988, S. 11–13).

Textbasierte Kommunikation dagegen – hier verstanden als medial schriftlich, unabhängig ihrer konzeptuellen Struktur – stellt demgegenüber eine eigene, zerdehnte Form der Interaktion dar, die unter besonderen Bedingungen stattfindet. „Zerdehnt“ meint in diesem Fall eine Verständi-

gung über Raum und Zeit hinausgehend, wobei die ursprünglich direkte Konfrontation der Gesprächspartner in zwei selbständige und zeitlich wie örtlich unabhängige Einheiten zerlegt wird. Der Text tritt als Medium zwischen die kommunizierenden Parteien (vgl. Becker-Mrotzek/Drommler 2006, S. 19). Der Leser nimmt dabei die Abwesenheit des Schreibers nicht nur gezwungenermaßen in Kauf, sondern befindet diese für gut, da in den wenigsten Fällen ein Leseakt Sinn macht, dessen Inhalt zur selben Zeit ebenso gut mündlich erfahrbar ist (Ausnahmen dazu sind genuine Fixierungen des mündlich abgesprochenen, wie bspw. im Vertragswesen) oder Protokolle, welche aber eher der erneuten Kommunikation zu einem späteren Zeitpunkt dienen bzw. als Referenz für externe, nicht an der direkten Kommunikation Beteiligte dienen. In der Konfrontation mit Geschriebenem liegt daher für gewöhnlich ein Moment der Asynchronität (vgl. Aust 1983, S. 2f), was eine andere Art der Perzeptionstheorie notwendig macht.

Interessant ist dabei besonders die Betrachtung der Genese der Kommunikationssituation. Wenn diese bei der direkten Kommunikation als recht eindeutig zu identifizieren ist, so stellt sich doch für das Lesen die Frage nach dem Kommunikationsanfang und ebendiesem -ende. Zu beantworten wäre diesbezüglich, ob sich die Kommunikation schon im Moment des Verfassens vollzieht oder erst während des Lesens, genauso wie, ob Schreiben und Lesen Teile einer einzigen Kommunikation oder zwei unabhängige und getrennte Ereignisse sind. Es lässt sich auf der einen Seite annehmen, dass der Schreiber eine Kommunikation beginnt, aber auf der anderen kann es ebenso sein, dass er sie durchaus schon fortsetzt oder gar beendet; vielleicht beendet aber auch der Leser durch seine Rezeption eine Kommunikation, setzt sie fort oder beginnt sie erst (vgl. Aust 1983, S. 2f). Daraus könnte man nach Aust den Schluss ziehen, dass Lesen immer in mehrere Kommunikationsakte einbezogen ist und erst von daher seine jeweilige Interpretation erhält.

Die Fixierung sprachlicher Äußerungen oder verbaler Gedanken in geschriebenen Texten erfordert andere Kompetenzen, nämlich oberflächlich die Fähigkeit zu schreiben und zu lesen und tiefergehend betrachtet, die nutzbringende Interpretation der zerdehnten Kommunikationssituation auf Basis des hervorgegangenen Materials.

Aust schreibt dazu, dass sich bei der Analyse von Leseprozessen nicht nur die Frage stellt, ob der Code des Lesers mit dem Code des Verfassers eine Schnittmenge bildet, sondern auch, ob der vom Leser rekonstruierte Verfasser-Code mit dem tatsächlichen Code des Verfassers homomorph ist (vgl. Aust 1983, S. 2). Kurz gesagt fragt sich der Leser an verschiedenen Stellen, ob er einerseits das Schriftstück überhaupt verstanden hat und ob er es andererseits auch richtig, d.h. wie vom Autor gewünscht, verstanden hat – ob also seine Interpretationsleistung den kommunikativen Erfordernissen genügt.

Die textbasierte Kommunikation bedingt also gegenüber dem direkten Gespräch auf der einen Seite Erschwernisse für Produzent und Rezipient, welche da wären: die durch die Zerdehnung der Kommunikationssituation wegfallenden für die mündliche Kommunikation aber kennzeich-

nenden wechselseitigen Aktivitäten wie Rückmeldungen, Nachfragen oder Bestätigungen, wobei diese eine wichtige Unterstützung bei der Realisierung kommunikativer Absichten leisten. Hier wird offenkundig, dass die mündliche Kommunikation primär eine dialogische Struktur aufweist, wohingegen die schriftliche eine eher monologische Form annimmt (vgl. Klicpera/Gasteiger-Klicpera 1995, S. 5; Becker-Mrotzek/Drommler 2006, S. 19). Äquivalente Vergewisserungen über die Korrektheit der eigenen Deutung könne so nicht im Vollzug dialogischer Wechselseitigkeit erreicht werden, sondern erfordern gesonderte Anstrengungen, so dass die Aktivitäten des Lesers vom bloßen Aufnehmen von Wörtern hin zu Inhalten über die Vergewisserung des Erfassten bis zu textphilologischen und exegetischen Verfahren reichen kann (vgl. Aust 1983, S. 3).

Lesen und Schreiben können daher als schwieriger als Sprechen und Zuhören bezeichnet werden, da hier mit einem abwesenden, im Beruf oftmals sogar unbekannten Gegenüber kommuniziert wird/werden soll. Becker-Mrotzek/Drommler (2006) führen weiterhin an, dass auch objektiv dieser Verlust ein Erschwernis der Kommunikation darstellt, da der situative Handlungskontext erst einmal hergestellt werden muss. Im Beruf ist an der Stelle allerdings zu betrachten inwieweit dieser durch den konkreten Kommunikationsanlass ggf. wieder gegeben sein kann.

Auf der anderen Seite finden sich aber Vereinfachungen, die offensichtlich überwiegen und zu einer stabilen Ausbildung der Schriftkultur führen. Als wichtigster Punkt ist sicherlich die zu-meist vorliegende Handlungsentlastung zu nennen. So können theoretisch Lesen und Schreiben ohne unmittelbaren Zeitdruck erfolgen, d.h., eine Äußerung kann beim Schreiben mehrfach überarbeitet, der Text kann als Ganzes geplant und auch wiederholt gelesen werden. Dadurch entfällt der für das Gespräch typische „Zugzwang“ Äußerungen gleichzeitig verstehen, planen und – mit einem erwartungsvollen Gegenüber – produzieren zu müssen (vgl. Klicpera/Gasteiger-Klicpera 1995, S. 6; Becker-Mrotzek/Drommler 2006, S. 19).

Jedoch kann durch die Verwendung eines Textes in einem beruflichen Handlungskontext eine andere Art von Handlungsdruck entstehen – nämlich der in der Situation möglichst schnell die benötigten (richtigen) Informationen aus dem Text zu entnehmen, in die Sachlage zu integrieren und umsetzen zu können. Hieran wird ersichtlich, dass Lesekompetenz gerade im Beruf eine höchst komplexe Kompetenz ist.

Eine Besonderheit von vielen Schriftsprachen, so auch der deutschen, ist das Vorliegen einer alphabetischen Schrift. Diese ist dadurch gekennzeichnet, dass sie eine abstrakte Symbolstruktur aufweist: die der Repräsentation von Phonemen²³ durch einzelne Schriftzeichen (graphemische Symbole). Diese sind ebenso arbiträr wie konventionalisiert²⁴, d.h. sie weisen keine analoge Beziehung zwischen Bezeichnendem und Bezeichnetem auf, was den Leseprozess insbesondere für ungeübte Leser bereits auf der Wortebene de facto zu einem Dechiffrierprozess

²³ kleinste bedeutungsunterscheidende akustische Einheit des Lautsystems einer Sprache

²⁴ vgl. dazu ausführlich Charles Peirces Theorien zur Semiotik u.a. 1906, 1977, 1986-1994 und 1993.

macht, und sind dennoch allgemein verständlich durch gesellschaftliche Übereinkünfte bezüglich Aussehen, Verwendung und Bedeutung. Weiterhin finden sich jedoch bspw. im Deutschen bei der schriftlichen Sprachverwendung auch zusätzlich gegenteilige Strömungen zur Verwendung einer Alphabetschrift. So sind Hinweistafeln in Bahnhöfen, Flughäfen und Massentransportmitteln durch eine formalisierte piktographische Struktur gekennzeichnet (vgl. Klicpera/Gasteiger-Klicpera 1995, S. 8). Die Gründe dafür liegen auf der Hand: zunehmende Internationalisierung und dem entgegenkommende relative Universalität analoger Repräsentationen.

Insgesamt lässt sich beobachten, dass auch und grade in der Arbeitswelt piktographische Darstellungen zum Alltag gehören. Wenn Informationen kurz, prägnant und möglichst allgemein verständlich sein sollen, wird häufig auch auf ikonographische Strukturen mit einer deutlich sichtbaren Beziehung zwischen Signifikant²⁵ und Signifikat²⁶ zurückgegriffen (bspw. bei bestimmten Hinweisen bzgl. des Arbeitsschutzes: statt eines Schildes mit der Aufschrift: „Hier muss/sollte ein Helm getragen werden!“ oder verkürzt: „Helm tragen!“ wird ein Schild mit einer abstrahierten Darstellung eines Menschen mit Helm gewählt). Sie unterscheiden sich noch von konkreten Abbildungen durch ihren Abstraktionsgrad.

Alles bisher Genannte bedarf trotzdem noch einer rezipientenseitigen Interpretation, um in einer speziellen Kommunikationssituation erfolgreich eingesetzt zu werden. Zusammenfassend festgehalten gibt es keine Art der selbsterklärenden textuellen Repräsentation von Kommunikation.

Damit diese Problematik möglichst gering ausfällt, etablierten sich vermutlich verschiedene Formen von Texten – sogenannte Textsorten. Dieser Betrachtung von Spezifikationen beruflich relevanter Textsorten widmet sich Kapitel 3.3.

Leseintention

Die Leseintention zu betrachten scheint aus verschiedenen Gründen lohnend. Sie ermöglichen einerseits eine trennscharfe Klassifikation der Leseprozesse und andererseits liegen bereits Hinweise darauf vor, dass die daran gekoppelten Verarbeitungsprozesse differieren. Diese scheinen sogar so stark von der Leserintention abzuhängen, dass in Experimenten unterschiedliche Instruktionen (zur Induzierung unterschiedlicher Leseintention) auch zu unterschiedlichen Verarbeitungsprozessen führten (Schmalhofer/Glavanov 1986). Schmalhofer/Glavanov haben dazu in ihrer Untersuchung 64 Studierende der University of Colorado, die keinerlei Wissen über das PC Programm LISP hatten, an einem Experiment teilnehmen lassen. Als Material wurde ihnen eine kurze Bedienungsanleitung (742 Wörter) präsentiert, welche in die LISP Datenrepräsentation („Atoms“ Einzelwerte und S-expressions Listen) einführt. Diese wurde basierend auf den ersten paar Seiten von McCarthys, Abrahams', Edwards, Harts und Levins LISP 1.5 Programmierhandbuch von 1965 konstruiert. Alle Probanden wurden angewiesen, die LISP

²⁵ (=Bezeichnendes) In der strukturalistischen Linguistik und Semiotik die Ausdrucksseite eines sprachlichen Zeichens, seine wahrnehmbare Form.

²⁶ (=Bezeichnetes) In der strukturalistischen Linguistik und Semiotik die Inhaltsseite eines Zeichens, seine nicht fassbare Form.

Anleitung in einer Weise zu bearbeiten, die es ihnen ermöglichen würde, den Test, der folgen würde, gut bis sehr gut zu bestehen. Der einzige Unterschied in der Instruktion der beiden betroffenen Gruppen der Studie war folgender: Der der Hälfte der Versuchspersonen wurde gesagt, dass der Test aus dem Schreiben einer kurze Zusammenfassung der LISP-Bedienungsanleitung bestehen würde. Der andere Hälfte wurde gesagt, dass der Test eine Programmieraufgabe, in Form des Schreibens und der Überprüfung der LISP Ausdrücke, sei (vgl. Schmalhofer/Glavanov 1986, S. 281f).

Festhalten ließ sich zunächst, dass die Probanden mit dem Leseziel „Textzusammenfassung“ die langsameren Leser waren. Insgesamt legen die Befunde nahe, dass die Textzusammenfassungsleser die makrostrukturellen Prozesse betonen und eher darum bemüht waren eine Textbasis zu konstruieren, während die anwendungsbezogenen Leser darauf fokussiert waren ein Situationsmodell zu entwickeln, indem sie mehr die substantiellen Informationen über LISP verarbeiteten. Sollten diese Ergebnisse tatsächlich die Produkte zweier unterschiedlicher kognitiver Prozesse oder Prozesskomponenten sein, so müsste sich dies nach Schmalhofer/Glavanov auch in anderen Resultaten zeigen (vgl. Schmalhofer/Glavanov 1986, S. 284). So zeigt sich in einem anschließenden Versuch, dass es eine signifikante Interaktion zwischen Leseziel und kognitiver Repräsentation gibt und dass Textzusammenfassungs-Leser eine bessere propositionale Textrepräsentation entwickeln, während sich anwendungsbezogene Leser auf die Konstruktion eines Situationsmodells konzentrieren. Auf Grund der Demonstration der Entwicklung der mentalen Struktur kann dies als zusätzlicher Beleg für die Unterscheidung zwischen propositionaler Textrepräsentation und einem Situationsmodell betrachtet werden (vgl. Schmalhofer/Glavanov 1986, S. 287). Wenn Leser einen Text (bspw. eine Bedienungsanleitung) durcharbeiten, können drei unterschiedliche und ineinandergreifende Prozesskomponenten unterschieden werden: Verarbeitungsprozesse bezogen auf die Wörter, auf die textuelle Bedeutung oder auf die hinter dem Text liegenden Sachverhalte (vgl. Schmalhofer/Glavanov 1986, S. 291). Dass diese von der Leseintention beeinflusst werden, kann auf Grund der vorliegenden Befunde zumindest angenommen werden.

Die angesprochenen Leseintentionen könnten sich besonders zur Abgrenzung und Klassifikation von theoretisch anzunehmenden unterschiedlichen Leseprozessen eignen. Eine systematische empirische Überprüfung hinsichtlich der möglichen Unterschiede der Modelle und deren Signifikanz stehen allerdings noch aus, wären aber dringend notwendig. Einerseits um ein tieferes Verständnis des Leseprozesses zu gewinnen, denn es fehlen bisher intentionsberücksichtigende Lesemodelle. Andererseits aber auch um gezielte Fördermaßnahmen entwickeln und anbieten zu können, die sowohl die Motivation der zu Fördernden erhalten, was eine wichtige Komponente bei der Leseförderung darstellt (vgl. Petsch et al. 2008; Ziegler/Gschwendtner 2010) wie auch eine transparente adressatenbezogene Nutzbarmachung der geförderten Kompetenz bieten.

Leseintentionen und daraus folgende Konsequenzen und Beobachtungen des Leseprozesses lassen sich wie folgt (siehe Tabelle 2) darstellen.

Tabelle 2 – Taxonomie der Leseintentionen

	Lesen, um zu handeln	Lesen, um eine Leseaufgabe zu erfüllen	Lesen, um zu lernen	Lesen, um zu entspannen
Intention	Bewältigung einer Aufgabe, die nicht notwendigerweise Lesen erfordert	Bewältigung einer Aufgabe, die notwendigerweise Lesen erfordert	Wissenserwerb	Erbauung
Daraus abgeleitetes Ziel/Funktion	Umsetzen/Nutzen	Beantworten	Behalten	Relaxation
Klassifikation	Funktionales Lesen	Funktionales Lesen	Epistemisches Lesen	Kontemplatives Lesen
Material	deskriptionale + depiktionale Texte, meist kurz oder nur Ausschnitte	deskriptionale + depiktionale Texte unterschiedlicher Länge	deskriptionale + depiktionale Texte; meist länger	meist fiktive, deskriptionale Texte unterschiedlicher Länge
Kontext	beruflicher/privater Natur	meist: schulisch-studierender Kontext oder Testkontext	meist: schulisch-studierender Kontext	Freizeitkontext
Lernen	Kein typischer Lernprozess aus Texten, aber nicht-intendiertes Lernen möglich	Kein typischer Lernprozess aus Texten, aber nicht-intendiertes Lernen möglich	Lernen als konkretes Handlungsziel	(nicht-)intendiertes Lernen möglich, je nach individuellem Anliegen
Textsorten	meist: Instruktionstexte, Fragmente anderer Textsorten	meist: vorgegebene situationell angepasste Texte	häufig: didaktisch aufbereitete Texte, wissenschaftliche Texte	Belletristik oder Instruktionstexte aus dem Freizeitbereich
Handlungsdruck	ja	ja	nein	nein

Verarbeitungstiefe	mentale Modellierung des Textnutzens notwendig	Propositionale Ebene oft ausreichend, MM möglich	Verarbeitungstiefe abhängig vom Lernziel; MM möglich	Verarbeitungstiefe abhängig vom Leserwunsch
Notwendigkeit	Lesen als optionaler (meist optimaler) Weg Lesen als Problemlösestrategie	Lesen als oktroyierter Weg; Lesen als konkrete Aufgabe	Lesen als optionaler (meist optimaler) Weg Lesen als Problemlösestrategie	Lesen als Selbstzweck Lesen als „Problemlösestrategie“ ²⁷
Erfolgskontrolle	Die Handlung wurde erfolgreich ausgeführt	Richtige Beantwortung der gestellten Frage	Abfrage von Wissen auf propositionaler Ebene (→ Behalten), ggf. Hinweise auf MM durch Transferleistung	Nur subjektive Aussagen zur Befriedigung des Entspannungsbedürfnisses möglich
Priorität	Lesen als Sekundär-Ziel → funktional: um... zu...	Lesen als Primär-Ziel	Lesen als Sekundär-Ziel → funktional: um... zu...	Lesen als Primär-Ziel

²⁷ Das „Problem“, insofern es als solches klassifizieren werden kann, ist in diesem konkreten Fall fehlende Unterhaltung bzw. der Wunsch nach Unterhaltung.

Erkennbar wird, dass die Leseintention zumindest theoretisch eines der trennschärfsten Kriterien zur Abgrenzung von Leseprozessen zu sein scheint. Eine empirische Verifikation der These steht jedoch noch aus.

Wie bereits in Kapitel 2.2 dargestellt wurde, ist das berufliche Lesen vornehmlich geprägt durch handlungsbezogene Leseintentionen, die durch die Ausrichtung auf eine vom Leseprozess selbst unabhängige berufliche Handlung gekennzeichnet sind. Auf Grund der in Kapitel 2.2.1 und 2.2.2 gezeigten beruflichen Leseanlässe scheint es angebracht der Frage nach den im Beruf ablaufenden textrezeptiven Verfahren und notwendigen Kompetenzvoraussetzungen eine detailliertere Betrachtung typischer Textsorten, welche besonders in dem oben aufgezeigten Kontext einer spezifischen Handlung auftreten (ohne dass das Lesen selbst die Handlung ist) und im Beruf häufig anzutreffen sind (bspw. Instruktionstexte) oder von ihrem Sprachduktus und Intention auf ein bestimmtes Klientel zielen (bspw. Fachtexte) vorzuschalten.

3.1.3. Analyse des spezifisch beruflichen Lesematerials

Bei der Analyse beruflichen Lesematerials können verschiedene Betrachtungsweisen gewählt werden. Auf Grund der Orientierung der vorliegenden Arbeit wird der Fokus im Folgenden auf zwei Aspekte gelegt – die Integration in eine Handlungsstruktur und die Berücksichtigung bzw. Verwendung verschiedener Repräsentationsformate. Zu Beginn werden kurz die bisher genannten Struktur- und Eigenschaftsmerkmale der möglichen Repräsentationsformate an Hand der Theorien von Schnotz systematisiert und in einen Nutzungskontext eingebettet. Anschließend werden kurz zwei typische Berufstextsorten aus der Perspektive der angewandten Linguistik betrachtet.

Rekapitulation und Präzisierung: Repräsentationsformate

Da im Verlauf der Arbeit immer die Trias „Symbol – Hybrid – Ikon (als Analogon)“ als Grundkategorien vorausgesetzt wird, werden diese drei verschiedenen Arten von kommunikativen Transportmodi hier nochmals ausführlich und unter besondere Berücksichtigung ihrer Einbettung in kommunikative Verwendungszusammenhänge erörtert.

Eine Unterscheidung lässt sich in zwei grundlegend heterogenen Zeichensystemen vornehmen – Symbolzeichen und ikonische Zeichen (vgl. Peirce 1906).

Symbolzeichen sind gekennzeichnet durch eine genuin arbiträre Struktur und erhalten ihre Assoziation mit dem Gegenstand vorrangig durch Konventionen – bspw. Schriftzeichen einer Alphabet-Schrift (s.o.). Ikonische Zeichen dagegen besitzen eine analoge Struktur. Sie sind eher mit dem bezeichneten Gegenstand durch Gleichartigkeit oder abstrakte strukturelle Berührungspunkte verknüpft. Diese beiden Systeme von Zeichen sind bereits nicht überlappungsfrei. Werden bspw. die frühen Piktogrammschriften betrachtet, zeigt sich die Entwicklung vom Ikon zum Symbol – so kann das ursprüngliche Ikon „Vogel“ im Laufe seiner Verwendungsgeschichte auch in andere Kontexte integriert werden und zunächst für „fliegen“ und schließlich für „oben“

oder „Luft“ stehen. Hier ist gut zu erkennen, dass ab einem bestimmten Punkt eine Konventionalisierung eintritt, ohne deren Kenntnis die Bedeutung nicht mehr entnommen werden kann. Hier ist also bereits die oben beschriebene Trias zu finden.

Ikonische Darstellungen werden dabei nochmals unterteilt in realistische Repräsentationen wie Bilder, die sich durch eine konkrete Form der Ikonizität auszeichnen, und am anderen Pol logische Repräsentationen wie Diagramme, die sich im Gegensatz dazu durch eine abstraktere Ikonizität auszeichnen. Konkrete Ikonizität bedeutet, u.a. dass bei realistischen Repräsentationen eine gegenständliche Form der strukturellen Übereinstimmung mit dem jeweiligen Begriff besteht, indem Höhe durch Höhe und Farbe durch Farbe dargestellt wird, allerdings eingeschränkt durch die Darstellungsmöglichkeiten im zweidimensionalen Raum. Bei logischen Repräsentationen scheinen zwar auf den ersten Blick keine Ähnlichkeit mit dem jeweiligen Dargestellten zu bestehen, bei genauerer Analyse sind sie jedoch mit ihm durch gewisse komplementäre Strukturmerkmale (sogenannte Analogierelation) verknüpft und bilden so das Bezeichnete auf Grund ebendieser strukturellen Gemeinsamkeiten ab, d.h. repräsentierte Aspekte und repräsentierende Aspekte können verschieden sein, so lange die Relationen zwischen den Aspekten innerhalb des Bildes gleich zu den Relationen innerhalb des Gegenstandes sind (vgl. Schnotz 2003, S. 578).

Ausgehend von diesen unterschiedlichen Zeichensystemen werden von Schnotz (2003) die beiden bereits beschriebenen verschiedene Arten der Repräsentation differenziert: deskriptionale Repräsentationen und depiktionale Repräsentationen.

Deskriptionale Repräsentationen beschreiben einen Sachverhalt mithilfe von Symbolen – bspw. durch kontinuierliche Texte in Büchern. Hierbei wird von „extrinsischen Repräsentationen“ (Schnotz 2003, S. 578) gesprochen. Depiktionale Repräsentationen dagegen weisen keine solchen expliziten Relationszeichen auf. Sie besitzen eher inhärente Strukturmerkmale, die mit gewissen Struktureigenschaften des darzustellenden Gegenstands übereinstimmen. Sie werden als „intrinsische Repräsentationen“ (Schnotz 2003, S. 578) bezeichnet (vgl. Schnotz 2003, S. 578). Sie können als vollständig bezüglich einer bestimmten Informationsart bezeichnet werden, denn bei der depiktionalen Darstellung eines Objekts wird zwangsläufig nicht nur seine Form, sondern auch seine Größe und Orientierung mit dargestellt. Deskriptionale Repräsentationen hingegen ermöglichen es ohne weiteres, die Form eines Objekts zu spezifizieren, ohne dabei auf seine Größe und Orientierung eingehen zu müssen (vgl. Schnotz 2003, S. 579).

Auffällig ist, dass sich die beiden Zeichensysteme in ihren Nutzungseigenschaften unterscheiden. Deskriptionale Repräsentationen sind verhältnismäßig abstrakt, wohingegen depiktionale Repräsentationen relativ konkret und spezifisch sind. Daraus folgt, dass deskriptionale Repräsentationen eine höhere Ausdrucksmacht als depiktionale Repräsentationen besitzen. Dies äußert sich unter anderem darin, dass durch deskriptionale Formate allgemeine Negationen sowie allgemeine Disjunktionen formulieren werden können. Depiktionale Repräsentationen demgegenüber können nur spezifische Negationen ausdrücken und Disjunktionen können, wenn über-

haupt sinnvoll, nicht durch eine einzelne Darstellung, sondern nur durch mehrere aneinander gereihte Bilder expliziert werden. Auf der anderen Seite sind depiktionale Repräsentationen meist besser geeignet, um Inferenzen²⁸ zu vollziehen, da die gesuchte Information verhältnismäßig einfach abgelesen werden kann (vgl. Schnotz 2003, S. 578f).

Diese ausführliche Hinführung dient der Grundlegung der folgenden Darstellungen bezüglich der Etablierten Formen von beruflich häufig auftretenden Textsorten. Bei diesen haben sich bestimmte Verwendungskontexte von Repräsentationsformaten etabliert, die an der jeweiligen Stelle berücksichtigt und analysiert werden.

Beispiel I – Instruktionstexte

Es gibt verschieden Texttypen, die Handeln anleiten oder anweisen. Für Instruktionstexte gibt es im Deutschen heterogene und vielfältige Bezeichnungen, die miteinander konkurrieren, bspw. Gebrauchsanweisung, Gebrauchsanleitung, Betriebsanweisung, Betriebsanleitung, Betriebsvorschrift, Bedienungsanleitung,..., ohne dass deutlich erkennbar ist, welches kategoriale Unterscheidungsmerkmale wären, die durch den Gebrauch unterschiedlicher Bezeichnungen repräsentiert werden. Daher werden die Begriffe im Folgenden synonym verwandt und unter der Bezeichnung Instruktionstexte bzw. Tätigkeitsleitende Texte subsummiert. Einzig davon abzugrenzen, aber in dieselbe Kategorie gehörend, sind Arzneimittelgebrauchsinformationen – besser bekannt als Packungsbeilagen. Diese werden am Ende des Kapitels kurz gesondert auf ihre Spezifika hin untersucht.

Für die Vorstellung der Bedienungsanleitung als Beispiel für tätigkeitsleitende Texte sprechen an dieser Stelle der allgemeine Bekanntheitsgrad dieses Texttyps und die intuitive Erfassbarkeit seiner Charakteristika auch für Nicht-Linguisten. Ferner wird in diesem der Zusammenhang zwischen einer Handlungsintention und der dazugehörenden Lesesituation besonders deutlich. Saile führt dazu aus: „Bei keiner anderen Textsorte greifen Sprache und Handlung enger ineinander als bei Anleitungstexten“ (Saile 1982, S. 134). Denn es ist davon auszugehen, dass sobald der Leser eine Information aufgenommen hat, er sie sofort in Handlungen umsetzt und keine mentale Konstruktion größerer Zusammenhänge anstrebt (vgl. Zieten 1988, S. 155f). Wobei eine Handlung verstanden wird als Zustandsveränderung, die von einem Agens kontrolliert wird (vgl. Saile 1982, S. 143). Anforderungen an Bedienungsanleitungen sind in der Regel von der Anwendungssituation unter starkem Handlungsdruck bestimmt²⁹, daher müssen sie schnell verfügbar, leicht verständlich und übersichtlich gestaltet sein. Denn anders als andere Texte zielen Bedienungsanleitungen direkt und unmittelbar auf die Praxis ab – ihre Besonderheit ist also die praktische Einbettung, welche eine Reihe von Konsequenzen für die Textstruktur hat. Die Lektüre von Anleitungen ist dabei eine lästige Notwendigkeit für ein anders, übergeordnetes Ziel (meist: der Gebrauch eines Geräts) (vgl. Ehlich 1994, S. 120f).

²⁸ Schlussfolgerung

²⁹ Bspw. zur Behebung eines Fehlers oder zur Einrichtung eines neuen Geräts – wenngleich andere Kontexte nicht ausgeschlossen werden können.

Für das Anleiten kann es vier unterschiedliche Möglichkeiten geben, die meist in der Praxis in Kombination existieren (vgl. Pelka 1982, S. 77f):

1. Anweisen durch außersprachliche Mittel, wie eine persönliche Demonstration unter der Verwendung von außersprachlichen Kommunikationsmitteln, wie bestimmten Gesten. Das setzt voraus, dass beide Kommunikanden sich in raumzeitlicher Koinzidenz befinden.
2. Anweisen durch außersprachliche Mittel, wie die Verwendung von depiktionalen Strukturen, d.h. zeichnerisch-symbolischen und bildlichen Kommunikationsmitteln, wie technischen Zeichnungen und Abbildungen. Hierbei können Anweisungsproduzent und -rezipient sich in raumzeitlicher Antikoinzidenz befinden, denn sofern eine Anweisung einmal fixiert ist, eröffnet sich ein Einsatz in vielen verschiedenen Anwendungssituationen. Allerdings setzt die Nutzung zeichnerisch-symbolischer Mittel im Anweisungsprozess voraus, dass die Kommunikationspartner über die notwendigen Kompetenzen verfügen sie zu verstehen.
3. Anweisen durch sprachliche Mittel, wie die natürliche Sprache. Dazu müssen die Kommunikationspartner zumindest in direktem Kontakt stehen (bspw. durch fernmündliche Gespräche).
4. Anweisen durch sprachliche Mittel im schriftlichen Gebrauch. Hierbei gelten die gleichen Vor- und Nachteile wie für außersprachliche Mittel (räumlich-zeitliche Trennung möglich, Fixierung des Inhalts, Flexibilität in verschiedenen Anwendungssituationen, Kompetenzerfordernisse...). Allerdings gehen zusätzlich eventuelle prosodische Nuancierungen verloren, da für diese keine graphischen Korrelate (wie Kommata oder Doppelpunkte) existieren (vgl. Solfeld 2010, S. 214).

Ein Problem bei der Analyse von Bedienungsanleitungen ist die auftretende Heterogenität bezüglich Umfang, Textaufbau und sprachlicher Gestaltung der jeweiligen Texte, die sich aus der Heterogenität der Produkte, des seitens des Produzenten angenommenen Vorwissens (betriebsinterne oder betriebsexterne, sowie fachinterne oder fachexterne Rezipienten), sowie unterschiedlicher Rezeptionssituationen und Kommunikationssituationen ergeben (vgl. Pelka 1982, S. 83; Hensel 1989, S. 144).

Grob untergliedern lassen sich nach Möhn drei Subtypen von Instruktionstexten (vgl. Möhn 1991, S. 197):

1. Unmittelbar an das einzelne konkrete Produkt gebundene Texte, welche zusammen mit dem Gegenstand beim Kauf erworben werden (bspw. Gebrauchs-, Bedienungsanleitungen, Beipackzettel...)
2. Ratgebertexte in Form von Büchern, Zeitungen, Zeitschriften (bspw. Rezepte, Bastelanleitungen, Hobbyanleitungen...)
3. Produktrelationierende Texte bezogen auf Kaufentscheidungen in Form von Heften, Zeitungen oder Zeitschriften (Testberichte,...)

Auch Juhl (1994) geht von drei Kategorien aus, untergliedert diese aber nach Anwendungssituation (vgl. Juhl 1994, S. 199):

1. Sofortanleitungen (bspw. Feuerlöscher,...)
2. Lernanleitungen (bspw. Textverarbeitung,...)
3. Nachschlageanleitungen (bspw. Programmiersprachen,...)

Vergleicht man die zentralen Parameter der beiden Kategorisierungen, so kann beobachtet werden, dass bei Juhls Einteilung in diesem Kontext die Intention des jeweiligen Lesers eine große Rolle einnimmt. Eine Nachschlageeinleitung ist demnach definiert durch den Leser, der sie zum Nachschlagen benutzt, ein anderer Leser verwendet dasselbe Dokument möglicherweise als Sofortanleitung. Damit scheint sie für die Betrachtung der Textebene, wie sie in dieser Arbeit erfolgt, eher ungeeignet zu sein, da sie keine stabile Klassifikation ermöglicht. Somit bietet sich die Klassifikation von Möhn an, allerdings bleibt in seinen Ausführungen unklar, ob diese angenommen kategorialen Unterschiede zwischen produktbegleitenden Texten, Ratgebertexten und produktrelationierenden Texten auf der Textebene existieren und falls ja, in welcher Form sie sich manifestieren.

Unabhängig von der Beipflichtung zu einer dieser formallinguistischen Klassifikationsformen lässt sich feststellen, dass die meisten Menschen insbesondere im heutigen Europa durch die zunehmende Komplizierung der Umwelt durch den technischen Fortschritt in zunehmendem Maße im Umgang mit technischen Produkten, Geräten und Maschinen der verschiedensten Art beinahe täglich auf die eine oder andere Weise angeleitet/angewiesen werden oder auf Anweisungen angewiesen sind. Beispiele hierfür sind: Verkehrstechnik (Lösen eines Fahrscheins, Öffnen einer Zugtür...), Bürotechnik (PC Nutzung, Bedienung eines Laminiergerät,...), medizinische Labortechnik (Umgang mit dem Mikroskop, Sterilisator,...), Haushaltstechnik (Unfallvermeidung, bei der Bedienung bestimmter Geräte, wie einem Mixer, Wahl und Dosierung des richtigen Waschmittels und/oder Weichspülers und/oder Entkalkers,...) sowie die sogenannte Freizeittechnik (bspw. Schlagbohrmaschine, Digitalkamera, Modellbausatz, Heimtrainer,...) (vgl. Pelka 1982, S. 75).

In der europäischen Ökonomie der Moderne hat die in der Industrialisierung begonnene Entwicklung eine neue Spitze erreicht. Die von der Arbeitsteilung bestimmte, weit fortgeschrittene Spezialisierung und Sektorisierung der Produktionstechnik, die sich gleichermaßen in einer Fülle differenzierter, verschiedenartiger aufeinander bezogener und/oder penetrierender technischer Fächer zeigt, wie auch in einer funktional hochgradig differenzierten Betriebsorganisation, hat dazu geführt, dass nur noch in den wenigsten Fällen die investiven Geräte und Maschinen dort hergestellt werden, wo diese zur Herstellung der konsumptiven Produkte benötigt werden (vgl. Pelka 1982, S. 76). Auch heute – ca. 30 Jahre später – ist dieser Trend weiterhin in der Industrie und Konsumtechnik zu erkennen.

Daraus resultiert, dass auf Rezipientenseite sowohl der Laie als auch der Fachmann in vielfältiger Weise auf (technische) Anleitungen angewiesen sind (vgl. Pelka 1982, S. 76) und diese nicht mehr nur in spezialisierten Bereichen Notwendigkeit besitzen. Bezüglich der kommunikativen Verwendung kann beobachtet werden, dass Bedienungsanleitungen meist dann konsultiert werden, wenn der Gebrauch bestimmter Geräte weitere Informationen oder Instruktionen erforderlich macht, d.h. wenn diese Informationen beim Anwender nicht quasi automatisch vorhanden sind bspw. vor der ersten Inbetriebnahme oder bei Störungen und Defekten – kurz: um ein akutes Problem zu lösen. Wie häufig die Anleitungen in Summa konsultiert werden, hängt einerseits von der Frequenz des Gebrauchs des Gegenstands und andererseits von den individuellen Behaltens- und Lernleistungen ab (vgl. Pelka 1982, S. 82). Bedienungsanleitungen können somit als Vermittler zwischen jemandem, der die Kompetenz besitzt eine Sache anzuleiten und jemandem der darauf angewiesen ist, die Anleitung zu bekommen, interpretiert werden (vgl. Pelka 1982, S. 98; Küster 1982, S. 107). Auch Solfeld (2010) benennt als primären Zweck einer Bedienungsanleitung die Befähigung der Leser zu gewissen Handlungen, die in Form einer mehr oder minder fixierten Sequenz zu realisieren sind, wobei die Quaestio³⁰ der Instruktionstexte die möglichst klare Vermittlung ebendieser Handlungsschritte ist (vgl. Solfeld 2010, S. 203). Dabei besteht das kommunikative Interesse der Hauptstrukturäußerungen in der Spezifikation der einzelnen Handlungsschritte, deren sequentielle Abfolge eines der Charakteristika einer Bedienungsanleitung sind (vgl. Borneto/Cortelazzo 1988, S. 208; Solfeld 2010, S. 209). Es lässt sich an dieser Stelle beobachten, dass je unbekannter dem Rezipienten der Referenzbereich ist, desto bereitwilliger wird er die dirigierende Funktion des jeweiligen Textes, seinen Aufforderungscharakter mit allen seinen Implikationen akzeptieren (vgl. Küster 1982, S. 111).

Ferner bemerkenswert an Bedienungsanleitungen ist außerdem, dass diese durch eine explizite Norm in Bezug auf Inhalt und Form bereits seit 1974 geregelt sind (ehemals DIN 8418 – heute: DIN EN IEC 62079).

Das rührt daher, dass Bedienungsanleitungen auch eine juristische Bedeutung beigemessen wird, denn die zu bedienende Sache kann in den Händen von Laien zu Gefahren für den Anwender oder Dritte werden. Es hat sich der Grundsatz entwickelt, dass die Eröffnung einer

³⁰ Textfrage, d.h. die Frage nach dem Ziel/Sinn und zugleich Inhalt des Textes

Gefahrenquelle die Verpflichtung begründet, die zur Abwehr der Gefahr erforderlichen Vorkehrungen zu treffen, sowie dass jeder der eine Gefahr verursacht hat, alles, was ihm den Umständen entsprechend zugemutet werden kann, zu tun hat, um ebendiese Gefahr abzuwenden (vgl. Fotheringham 1982, S. 69f). Die Bedienungsanleitung dient dem Nachkommen dieser auferlegten Gefahrenabwehrpflicht, in dem hier vor den spezifischen Gefahren gewarnt und durch die Instruktion zu sachgerechtem Umgang angehalten wird, um dieses Ziel zu erfüllen, muss die Instruktion deutlich, ausreichend genau und umfassend sein (vgl. Fotheringham 1982, S. 70). Doch können nicht nur Konstruktions- und Fabrikationsmängel, sondern auch (sprachliche) Instruktionsmängel Gefahrenursache sein (vgl. Saile 1982, S. 151f).

Anleitungen haben gemäß dieser Norm als sogenannte produktbegleitende Texte die Aufgabe, dem Verwender die einschlägigen Kenntnisse zu vermitteln, die zum sachgerechten und sicheren Gebrauch oder Betreiben technischer Erzeugnisse erforderlich sind. Dabei umfasst der Gebrauch auch Instandhaltung, Wartung und Entsorgung, wozu im Detail Identifizierung, Produktspezifikation, Vorbereitung des Produkts für den Gebrauch, Grundlegende Funktionen, Sekundäre Funktionen, Optionale Module und Extras, Schutz von Personen, Instandhaltung und Reinigung, Informationen zu Sicherheit und Gesundheit sowie Widerspruchsfreiheit der Informationen gehören (vgl. DIN EN IEC 62079).

Auffällig ist, dass Bedienungsanleitungen nur in den seltensten Fällen reine Instruktionstexte sind. Sie sind vielmehr auch dadurch charakterisiert, dass sie auf illokutionärer³¹ Ebene neben instruktiven auch beschreibende, erläuternde, anleitende, vorschreibende und empfehlende Textteile enthalten (vgl. Pelka 1982, S. 80f; Fuchshuber-Weiß 1985, S. 9). Interessant dabei ist, dass dies bedeutet, dass der Rezipient zunächst den für seine konkrete Situation relevanten Inhalt innerhalb des Dokuments ermitteln muss, dessen Struktur zumindest einordnen können muss und dann bei Bedarf nur diesen wie gewünscht rezipiert (selektives Lesen).

Unter dem Aspekt der Verständlichkeit generell wird die Unterstützung des Textes durch bildliche Darstellungen als nützlich bezeichnet (vgl. Pelka 1982, S. 82).

Es herrscht in der einschlägigen Literatur (siehe unten) große Einigkeit bezüglich der Position, dass bei schriftlichen Anleitungssituationen, wie sie heute vermehrt auftreten, aktionale und gestische Mittel zumeist fehlen, wie auch der für mündliche Anleitungssituationen charakteristische Verständigungsdiskurs. Sowohl kennt der Autor einer Instruktion kaum seine Adressaten, ihre Fähigkeiten und Defizite, andererseits haben die unterschiedlichen Adressaten nicht die Option den Verfasser bei der Textproduktion zu unterbrechen oder ihm Gegenüber explizite Verbalsierungen des Nicht-Verstehens zu äußern. Daher werden bei Bedienungsanleitungen als Kompensation für das Fehlen deiktischer³² sprachlicher Verfahren Mittel des semiotischen Zeigens eingesetzt. Wird dieses versäumt, hat das dieselben Folgen wie das Weglassen oder Nicht-

³¹ d.h. der Sprechakt im Hinblick auf seine kommunikative Funktion

³² Demonstrativer/hinweisender...

Sehen von Zeigegesten in mündlichen Anleitungssituationen – Beziehungen können nicht hergestellt werden (vgl. Zieten 1988, S. 257; Borneto/Cortelazzo 1988, S. 209; Hartung 1992, S. 123; Fickermann 1994, S. 80; 88; Ehlich 1994, S. 126; 131).

Pelka (1982) liefert für die Beobachtung, dass zumeist die schriftsprachlich realisierten Anweisungen/Anleitungen in Kombination mit zeichnerisch-symbolischen und bildlichen Darstellungen dominieren, verschiedene Gründe:

Die schriftliche Fixierung von Wissen, welches für eine sichere und sachgerechte Handhabung erforderlich ist, ermöglicht jedem jederzeit auf rationelle Weise Zugang zu ebendiesem Wissen. Des Weiteren sind technische Erzeugnisse derzeit allgemein so komplex, dass ihre Handhabung, d.h. Bedienung ohne exakte Instruktion kaum mehr möglich ist. Ferner werden technische Erzeugnisse heute fast ausnahmslos in serieller Massenproduktion gefertigt, dabei ermöglicht eine fixierte Anleitung einheitliche und verbindliche Information. Sodann bestehen heute bezüglich der Nutzung technischer Geräte und Maschinen justiziell geregelte Vorschriften betreffend der Sicherheit und des Schutzes der Anwender, die vom Produzenten genauestens zu beachten sind und über die auch der Nutzer exakt instruiert werden muss (s.o.). Abschließend ist ein weiterer Grund die heute existierende Print- und Distributionstechnologie, welche eine massenhafte Herstellung und Verbreitung von Instruktionen in typographisch angemessener und repräsentativer Form ermöglicht (vgl. Pelka 1982, S. 78f). Daneben können durch die Benutzung von Bildern im Text vorhandene Vagheiten und Mehrdeutigkeiten eingeschränkt werden (vgl. Saile 1982, S. 137). Koch hält daher zum Thema Abbildungen innerhalb von Instruktionstexten sogar fest: „...good instructional texts should show a high degree of iconicity“³³ (Koch 2008 zitiert nach Solfeld 2010, S. 203).

Damit gibt es bei Instruktionstexten zumeist drei gegenseitig bedingte Komponenten: Text – Bild – Situation.

Saile entwickelte dazu folgende Graphik:

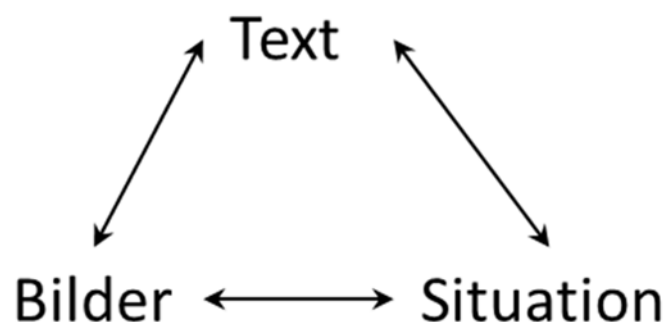


Abbildung 3 – Interpretationskomponenten der Kommunikation mit Instruktionstexten (Saile 1982. S. 135)

³³ In etwa: „... gelungene Instruktionstexte sollten einen hohen Grad an Ikonizität besitzen.“

Alle drei Komponenten überlappen sich gegenseitig ohne dabei identisch zu sein. So können bspw. in Bildern Dinge gezeigt werden, die im Text nur angedeutet werden, andererseits können aber im Text Abläufe beschrieben werden, die in den Bildern zusätzlich unterstützt werden oder in der Situation weitere Aspekte zum Tragen kommen, die keine repräsentative Entsprechung haben (vgl. Saile 1982, S. 141). Daher bietet sich folgende alternative Darstellung an:

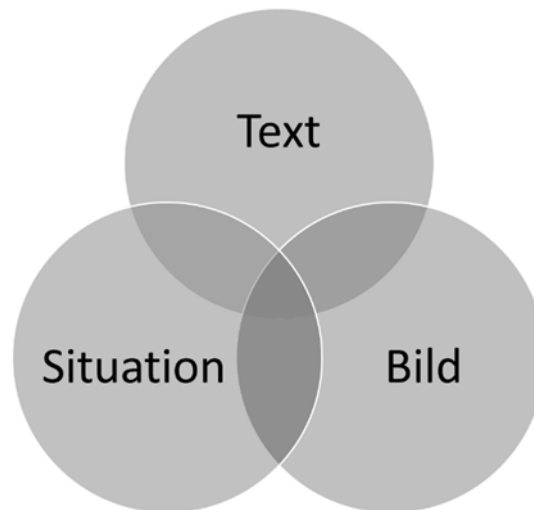


Abbildung 4 – Weiterentwicklung der Interpretationskomponenten der Kommunikation mit Instruktionstexten

Begründungen für die Schwierigkeiten im Umgang mit Bedienungsanleitungen können in der Heterogenität der Benutzergruppen, deren unterschiedliches Vorwissen und unterschiedliche Fragestellungen/Interessen, gefunden werden. Auf der konkreten Textebene lassen sich unübersichtliche Textgliederung, Kompliziertheit technischer Zeichnungen und Bilder, nicht eindeutige Zuordnung von depiktionalen und deskriptionalen Stimulus und zu komplizierte Fachsprache insbesondere die Verwendung unverständlicher Fachtermini nennen (vgl. Pelka 1982, S. 83). Solfeld führt auch die Abhängigkeit der Qualität einer Bedienungsanleitung von der (fehlenden) Ikonizität an (vgl. Solfeld 2010, S. 222f).

Bei Bedienungsanleitungen lässt sich wie bei keiner anderen Textsorte das Gelingen der Kommunikation an Hand der Frage „Wurde die zugrundeliegende Handlung erfolgreich durchgeführt?“ überprüfen. Dies ist über einen Ausführungstest ermittelbar. Dabei können die Zeit, die für die Ausführung der verlangten Aufgabe benötigt wird, sowie die bei der Ausführung gemachten Fehler Anhaltspunkte für die Verstehensleistung geben. Voraussetzung ist, dass Intelligenz, technische Erfahrung/Vorwissen und manuelle Geschicklichkeit der Benutzer durch entsprechende weitere Tests kontrolliert wird (vgl. Saile 1982, S. 150). Saile bleibt allerdings auf der theoretischen Ebene und trägt seine Überlegungen nicht in die Praxis. Auch Ehlich verweist auf den Bedarf von empirischen Untersuchungen über den Umgang der Leser mit Instruktionstexten (vgl. Ehlich 1994, S. 136). Überraschenderweise finden sich jedoch bisher keine solchen Tests als Indikatoren für die Lesekompetenz im Handlungskontext. Dies sollte dringend in Angriff genommen werden, da sich hieraus wichtige Hinweise für Leseprozesse und auch Intentionseinflüsse gewinnen lassen.

Abschließend lassen sich als weitere allgemeine Textcharakteristika einige Auffälligkeiten in aller Kürze festhalten:

Die Textcharakteristika in Bedienungsanleitungen geben einen expliziten Hinweis auf die Textfunktion (vgl. Pelka 1982, S. 86), denn die Texte bestehen immer in einem gewissen Maß aus sprachlichen Anweisungen eine Handlung zu tun oder zu unterlassen (vgl. Pelka 1982, S. 96). Charakteristisch ist dabei der kombinierte Gebrauch sprachlicher und außersprachlicher Kommunikationsmittel (vgl. Pelka 1982, S. 93). Bemerkenswert ist die Funktionalisierung der depiktionalen Repräsentationen: Es existieren bspw. viele Abbildungen konkreter Handlungen und Details bezüglich des Produkts, jedoch keine von Personen, die die Bedienungsanleitung lesen (vgl. Pelka 1982, S. 89). Verglichen mit anderen Textsorten scheinen Bedienungsanleitungen relativ durchgegliedert zu sein, d.h. es gibt sprachliche Anzeigen von Textabschnitten durch z.B. auffällige Überschriften und/oder in Inhaltsverzeichnisse, um einen punktuellen Zugriff zu erleichtern. Auch sind Hinweise, die besonders beachtet werden sollen, häufig räumlich von den umgebenden Textpassagen getrennt und (zusätzlich) durch sprachliche und/oder typographische Mittel hervorgehoben (vgl. Pelka 1982, S. 91f). Fachausdrücke bleiben interessanterweise weitestgehend unerklärt (vgl. Pelka 1982, S. 95). Insgesamt scheint ein sprachökonomischen Primat bei der Verfassung bemerkbar zu sein (vgl. Pelka 1982, S. 96).

Die Erstellung dieser Dokumente liegt in der Regel bei den jeweiligen Produktherstellern und erfolgt vielfach in betriebsorganisatorisch eigenständigen Abteilungen, in den betriebsinternen Werbeabteilungen oder bei außerbetrieblichen Agenturen, die nach Angaben von beauftragenden Firmen die Dokumente erstellen (vgl. Pelka 1982, S. 80). Es hat sich so eine neue Zulieferindustrie entwickelt, die vornehmlich Anleitungen produziert oder ihren Produktionsprozess begleitet (vgl. Möhn 1991, S. 187).

Somit können technische Instruktionstexte als exemplarisch für zwei entscheidende Faktoren gesehen werden: es sind authentische Instruktionstexte, die sowohl auf einen konkreten Handlungsrahmen referenzieren, als auch als exemplarisch für hybride Repräsentationsformate gesehen werden können. Eine weitere Betrachtung und Einbindung in Assessments scheint daher lohnend.

Arzneimittelgebrauchsinformationen weisen viele Parallelen zu Bedienungsanleitungen auf. So sind auch sie Handlungsanleitungen, die in diesem Fall sogar vom Gesetzgeber im Arzneimittelgesetz (§11 – Gestaltung und §12 – Warnhinweise) explizit hinsichtlich Form und Inhalt normiert wurden (vgl. Hoffmann 1983, S. 138). Der Adressatenkreis ist auch hier bezüglich seiner Anforderungen, Lesegewohnheiten und Vorkenntnisse als heterogen einzustufen (vgl. Gloning 1995, S. 45). Der an dieser Stelle relevanteste Unterschied zu Bedienungsanleitungen liegt eher in der Verwendung der Repräsentationsformate: Bedienungsanleitungen sind oftmals, wie oben dargestellt, Hybridformen mit großem depiktionalem Anteil. Arzneimittelgebrauchsinformationen wurden zwar bisher nicht unter diesem Aspekt untersucht, dennoch lässt sich festhalten, dass in einschlägigen Betrachtungen (Hoffmann 1983; Gloning 1995) keine depiktionalen Strukturen erwähnt werden und eine nicht repräsentative eigene Untersuchung

auch keine derartigen Hinweise fand, so dass die Aussage, dass Packungsbeilagen vornehmlich deskriptionale Repräsentationen diskontinuierlicher Form sind, als zutreffend angenommen werden kann.

Sodann können auch Arzneimittelgebrauchsinformationen als interessant für Assessments gesehen werden, da es ebenfalls authentische Instruktionstexte sind, die einen konkreten Handlungsrahmen referenzieren, jedoch ein anderes Repräsentationsformat aufweisen und so die Möglichkeit eines Vergleichs eröffnen.

Beispiel II – Fachtexte

Eine andere Art von Texten, die in beruflichen Kontexten verstärkt im Bereich der schulischen Ausbildung auftreten sind sogenannte Fachtexte.

Die Relevanz dieser Textgruppe wird offenkundig, wenn bspw. die Ausführungen des deutschen PISA-Konsortiums betrachtet werden: Entscheidend für eine erfolgreiche Bildungskarriere von Kindern aus Migrantenfamilien sei die Beherrschung der deutschen Sprache auf einem dem jeweiligen Bildungsgang angemessenen Niveau (vgl. Baumert et al. 2001a, S. 374). Unfraglich scheint dabei, dass dies auch für deutschsprachige Lernende zutrifft. Als besonders wichtig wird dabei die Fähigkeit, Texte und hier speziell Fachtexte verstehen zu können, hervorgehoben, da das Verstehen von Fachtexten eine Bedingung für die Auseinandersetzung mit Fachinhalten sowie für den Erwerb und den Ausbau von Fachwissen und Fachkompetenzen ist (vgl. Fluck 1992, S. 142; Oldenburg 1992, S. 38).

Die für berufliche Lernprozesse im Besonderen relevante Fachtextsorte ist die der Lehrtexte. Die Vorbereitung eines beruflichen Lebens kann ohne die Beschäftigung mit diesen Texten nur schwer gelingen, denn das Bestehen von Abschlussprüfungen im Rahmen der Berufsausbildung ist ohne ausreichendes Leseverstehen von Fachtexten kaum möglich (vgl. Niederhaus 2011, S. 209). Auch wenn im direkten Berufsalltag die Konfrontation mit diesen Texten eher gering ausfällt, so stellen sie doch den Zugang – respektive die Zugangsbeschränkung – für den Erwerb des Berufs und den anschließenden beruflichen Aufstieg dar, wohl meist, ohne, dass es den Aspiranten derart bewusst ist.

Fachtexte können sowohl Mittel als auch Ergebnis fachlicher Kommunikationsbeziehungen sein (vgl. Hoffmann 1985, S. 233f). Sie haben die Funktion, eine präzise, effektive und situativ adäquate Kommunikation über fachliche Sujets zu gewährleisten (vgl. Gläser 1990, S. 6). Sie werden definiert als

„Instrument und Resultat der im Zusammenhang mit einer spezialisierten gesellschaftlich-produktiven Tätigkeit ausgeübten sprachlich-kommunikativen Tätigkeit. [Sie bestehen] aus einer endlichen, geordneten Menge logisch, semantisch und syntaktisch kohärenter Sätze, oder satzwertiger Einheiten, die als komplexe sprachliche Zeichen komplexen Aussagen im Bewusstsein des Menschen und komplexen Sachverhalten in der objektiven Realität entsprechen.“ (Roelcke 2005, S. 21)

Dabei kennzeichnet sie das Vielfache Vorkommen bestimmter sprachlicher Strukturen auf fast allen sprachlichen Ebenen.

Auf der Textebene ist für Fachtexte bspw. die Nutzung außersprachlicher bzw. spracharmer Kommunikationsverfahren bzw. das häufige Vorkommen von depiktionalen Repräsentationen typisch. Dazu zählen einerseits informierende Bilder wie Abbilder (z.B. Fotos, Zeichnungen, grafische Symbole), logische Bilder (v. a. Diagramme) und schematische Bilder (z.B. Schaltpläne, Landkarten). (vgl. Buhlmann/Fearns 2000, S. 64; Roelcke 2005, S. 92f; Weidenmann 1994, S. 9).

Ein weiteres Merkmal von Fachtexten ist, dass sie auf der Satzebene dadurch gekennzeichnet sind, dass die Sätze länger sind als in anderen Registern (Hoffmann 1987, S. 205). Überdies charakterisiert sie eine Art Textkompression oder sprachliche Ökonomie vor allem durch Konstruktionen, welche der Einsparung von Nebensätzen dienen – bspw. durch präpositionale Wortgruppen oder Partizipien im attributiven Gebrauch. Die Präzision als weiteres Kennzeichen von Sachtexten wird sprachlich durch die gehäufte Verwendung von Attributen erreicht. Um dem gewünschten Wahrheitsanspruch Gültigkeit zu verleihen, die Objektivität zu betonen und die Allgemeingültigkeit fachbezogener Aussagen zu unterstützen, wird in Fachtexten zudem häufig ein anonymer Sprachstil verwendet bzw. eine passivische, abstrakte Ausdrucksweise benutzt (vgl. Oksaar 1998, S. 397).

Auf der Wortebene schließlich sind zahlreiche Wortschöpfungen eine Typik von Fachtexten. Diese dienen beständig als Erweiterung, Ergänzung und Ausdifferenzierung des Fachwortschatzes. Zu den wesentlichen und häufigen Wortbildungen in Fachtexten gehören bspw. Komposita, Derivate, Fremdwörter und Abkürzungen (vgl. Buhlmann/Fearns 2000; Fraas 1998; Roelcke 2005).

Aus Sicht der Pragmatik gibt es eine weitere bemerkenswerte Eigenschaft von Fachtexten. Diese liegt in ihrer auf den Rezipienten gerichteten Intentionalität. Sie sind ausgerichtet quasi komponiert auf einen bestimmten Zweck im Rahmen der Kommunikation innerhalb eines bestimmten Tätigkeitsbereiches. Der fachliche Textproduzent verfolgt mit dem betreffenden Text das Ziel, einem Rezipienten bestimmte Gegenstände und Sachverhalte nahezubringen. Sie sind also ähnlich wie Bedienungsanleitungen nicht als autorexpressive Formen konzipiert, sondern als rezipientenfokussierte.

Interessant ist die Frage, die Niederhaus (2011) stellt: Wie gestaltet sich die Ausprägung an Fachtextlichkeit in Lehrbüchern verschiedener Berufsfelder? Dazu wird weiter ausgeführt:

„Denkbar ist einerseits, dass Lehrbuchtexte verschiedener Berufsfelder einen ähnlichen Fachsprachlichkeitsgrad aufweisen, da es sich zwar um verschiedene Fächer handelt, die Textsorte jedoch die gleiche ist. Es kann andererseits angenommen werden, dass sich Lehrbuchtexte in Abhängigkeit ihres Faches unterscheiden, dass also das Fach den Grad der Fachsprachlichkeit von Texten bestimmt.“ (vgl. Niederhaus 2011, S. 213).

Sie kommt zu dem bemerkenswerten Ergebnis, dass sich die Fachkundebüchern verschiedener Berufsfelder linguistisch signifikant voneinander unterscheiden, was als Argument für eine integrierte fachspezifische Sprachförderung in der beruflichen Bildung, also für eine enge(re) Verzahnung von Fach- und Fachsprachenlernen, dienen kann. Ferner zeigt sie ein, insbesondere für die Berufspädagogik, interessantes Feld interdisziplinärer Forschung auf, nämlich die systematische Untersuchung des Zusammenhangs zwischen Fachsprachlichkeitsgrad von Fachkundebuchtexten und dem Grad der Fachlichkeit eines Berufsfelds insgesamt (vgl. Niederhaus 2011, S. 221f).

Beispiel III – normative Texte, Texte des Rechtswesens und der Verwaltung

Üblicherweise erfolgt die Klassifikation von Textsorten gemäß verschiedener Merkmale. In verschiedenen Ansätzen wird das Kriterium der Textfunktion als das entscheidende bezeichnet – hinzukommen u.a. Kohärenzstrukturarten, sprachliche Gestaltung (syntaktische und lexikalische Mittel), thematische Struktur, dominante Texthandlungen etc. (vgl. Busse 2000, S. 658)

Bezüglich der „Textsorte“ juristische Texte bzw. hinsichtlich ihrer Differenzierung, existiert jedoch an dieser Stelle folgendes von Busse (2000) ausgeführtes Problem,

„dass die hierfür in nahezu allen Klassifikationsansätzen für Textfunktionen vorgesehene Kategorie ‚normative Textfunktion‘ zumindest aus der Binnensicht einer sich auf den Gegenstand und seine Spezifika wirklich einlassenden rechtslinguistischen Forschung als Restkategorie erscheint, in die pauschal alles hineingepackt wird, was irgendwie dem Bereich Recht zugeordnet werden kann; interne Differenzierungen wurden bisher üblicherweise von Linguisten nicht vorgenommen, ebenso wenig wie eine tiefere Analyse der Spezifika juristischer Texte zu beobachten ist.“ (Busse 2000, S. 658)

Busse deutet an, dass seiner Meinung nach, eine Klasse von „normativen Texten“ viel zu weit gegriffen wäre, um eine homogene Textsorte darzustellen und dass somit auch eine Beschreibung dieser vermeintlich einheitlichen Textsorte den spezifischen Besonderheiten der jeweiligen Texten und ihren Funktionen nicht gerecht wird. Dennoch soll an dieser Stelle versucht werden, eine derartige Analyse durchzuführen und zumindest globale Merkmale feststellen zu können, damit in etwa festgehalten werden kann, welche Schwierigkeiten beim Lesen von normativen Texten auftreten können und evtl. welcher Repräsentationsformate sich bedient wird.

Festhalten lässt sich nach Busse (2000) trotz der genannten Schwierigkeiten, dass normative Texte Schriften sind, welche von innen oder von außen auf eine festgefügte und eigenständige gesellschaftliche Institution – hier gemeint: das Recht – bezogen sind und damit auf einen deutlich von anderen zentralen Textklassen (etwa denen des Alltagslebens) abgrenzbaren Bereich verweisen (vgl. Busse 2000, S. 658). Auf Grund der wichtigen Position, die das Recht in vielen gesellschaftlichen Institutionen und Handlungsbereichen besitzt, weisen auch andere, nicht primär der Jurisdiktion zuzurechnende Texte, Merkmale rechtlicher Texte auf.

Dabei fällt auf, dass auch Textsorten in anderen rechtsnahen Bereichen quasi juristische Merkmale so dann auch Merkmale juristischer Textsorten im weitesten Sinne aufweisen, ohne damit

jedoch als genuine Texte des Rechtswesens und der Justiz zu sein. Im Besonderen trifft dies auf die Textsorten (staatlicher) Verwaltung zu, welche auf Grund der im deutschen Rechtssystem und dem deutschen Staatsaufbau geltenden Rechtsförmigkeit³⁴ der Verwaltung bisweilen größere Überschneidungen zu den konkreten Textsorten des Rechtswesens (z.B. Beschluss, Bescheid, Verordnung, Verwaltungsakt) besitzen (vgl. Busse 2000, S. 622). Damit wird deutlich, dass eine lediglich sprachbezogene oder verwendungsbezogene Abgrenzung von Textsorten hier nur schwer erfolgen kann, stattdessen ist eine Berücksichtigung der funktionaler Merkmale notwendig, welche sich u.a. aus den institutionellen Regeln, Verfahrensweisen und wechselseitigen Abgrenzungen abgeleitet werden können.

Auch eine Reduktion auf die vermeintlich eindeutigeren und besser eingrenzbarer Klasse der normativen Texte ist hier nicht zielführend, da auf Grund der Komplexität des juristischen Normbegriffs derartige Texte ebenso schwer als eigene Textklasse abgrenzbar sind. Primär würden darunter Gesetze (einschl. Verfassungstexte und Verordnungen) fallen, jedoch wirken normativ im weiteren Sinne auch Urteilstexte (wenn sie von höheren Instanzen stammen) und rechtswissenschaftliche Texte (wie Kommentartexte) – die Normativität erstreckt sich damit quasi über mehrere Textklassen hinweg (vgl. Busse 2000, S. 663). Ferner würden bei dieser Art der Einteilung jedoch die gesamte Klasse von Verwaltungstexten ignoriert werden, welche keinen normativen Charakter haben, jedoch intuitiv bei der Betrachtung einzuschließen sind.

Eine der wichtigsten Besonderheiten von Rechtstexten ist, dass diese Textsorten Elemente institutionellen Handelns darstellen (vgl. Busse 2000, S. 663). Dies mag profan erscheinen, ist jedoch einer der wichtigsten Aspekte bei der Untersuchung der Anforderungen, die juristische Texte an den jeweiligen Leser stellen. Durch die Verfassung eines Dokuments und seiner Veröffentlichung bzw. Übermittlung an den jeweiligen Empfänger kann Realität geschaffen werden und eine Handlung ausgeführt werden. Damit, dass ein Gesetz verabschiedet wird, kommt es zu Folgen in der Realität ohne dass konkret weiteres Zutun notwendig ist. Sogar durch die Unterlassung eines schriftlichen Aktes kann eine Folge eintreten – bspw. die nicht Verfassung eines Widerspruchs. Wenn der Leser diese oder ähnliche Erfahrung bereits im Zusammenhang mit schriftlichem Verwaltungshandeln gemacht hat, kann dies seine Einstellung (und damit auch seine Herangehensweise an) bezüglich dieser Schriftstücke bestimmen. Texte im Bereich des Rechtswesens können daher zu seiner motivational ungünstigen respekts- und erwarteten Überforderungshaltung führen. Aus den genannten möglichen Folgen von Rechtstexten ergibt sich aber auch, dass an diese Schriften spezielle Anforderungen gestellt werden, welche sich auf Textebene in bestimmten Charakteristika niederschlagen.

³⁴ Zur Präzisierung: Rechtsförmigkeit Sinne erfasst außer der konkreten staatlichen Verwaltung weitere eher als staatsfern zu bezeichnende gesellschaftliche Bereiche, wie Vereine, Parteien, Kapitalgesellschaften, große Wirtschaftsunternehmen etc. Entscheidend ist, dass die innerorganisatorischen Rechtsverhältnisse einer förmlichen Bestätigung durch Institutionen des Rechtswesens und der Justiz bedürfen. Bsp.: Satzung – die einer Partei, eines Vereins oder einer Genossenschaft ist im Kern keine klassisch normative Textsorte oder eine des Rechtswesens, dennoch kann sie als juristische Textsorte im weiteren Sinne betrachtet werden, da sie (u.a. in den linguistisch beschreibbaren Merkmalen wie Textaufbau, Gliederungsprinzipien, Kohärenzstrukturen) viele Eigenschaften aufweist, die auch für juristische Texte festgehalten werden und die vermutlich grade aus diesem Grund aus deren Mustervorrat für sie übernommen wurden (vgl. Busse 2000, S. 662).

Wagner (1984) hält als Zweck von Verwaltungssprache fest: „Die Verwaltungssprache wird im Dienst von Beamten und Verwaltungsangestellten zur Erfüllung staatlicher Aufgaben gebraucht. Sie hat deshalb einen offiziellen, unpersönlichen und funktionalen Charakter.“ (Wagner 1984, S. 97)

Zu den oben erwähnten Charakteristika, die diesem Zweck dienen sollen, gehören grammatische Konstruktionen, welche sonst vor allem in wissenschaftlich-technischen Fachsprachen zu finden sind. Diese zielen im Kern auf vor allem zwei Dinge ab: eine sachliche Darstellung der Gegebenheiten und Ausdrucksökonomie. Die vermeintliche Versachlichung wird durch viele Nominalisierungen in Kombination mit bedeutungsarmen Verben erzeugt, die Ausdrucksökonomie zeigt sich hingegen an einer komprimierten Beschreibung von Inhalten durch Partizipal- und Infinitivkonstruktionen. Weiterhin kann eine vielfache Verwendung passivischer Konstruktionen unter Vermeidung der Nennung des Agens³⁵ beobachtet werden, was zur objektiven und sachlichen Darstellung beitragen soll. Als Tempus tritt in den meisten Fällen das Präsens auf, um so den Anspruch auf Allgemeingültigkeit der Aussagen zu unterstreichen (vgl. Hansen-Schirra/Neumann 2004, S. 169f).

Festzuhalten ist, dass diese Art der Sprache eine genuine Schreib- und Lesesprache ist, so dass alle Sätze (theoretisch) vollständig und unabhängig von einer spezifischen Situation oder einem konkreten Gegenüber verständlich sind. Ferner erlaubt die schriftliche Form eine weiter gespannte Übersicht und langsame Aufnahme zusammengedrängter Aussagen, so dass komplexere Konstruktionen, die den obengenannten Ausdruckszielen dienen, präferiert werden (vgl. Wagner 1984, S. 98).

Der erfahrungsgemäß häufig als unanschaulich und abstrakt kritisierte Stil des „Verwaltungsdeutsch“ rührt vermutlich aus der Unterordnung konkreter Einzelfälle unter allgemeingültige begriffliche Kategorien her, wobei substantivische Begriffe es gestatten die Sachverhalte exakt in bestimmte Rubriken, Register und Karteien einzuordnen (vgl. Wagner 1984, S. 99).

Eine weitere Schwierigkeit mit Rechts- und Verwaltungstexten ist wiederum eine parallel zu wissenschaftlichen Fachsprachen: Es wird eine große Menge von Fachwörtern und Bezeichnungen verwendet, die entweder Laien unbekannt sind oder eine andere Bedeutung besitzen als ihre umgangssprachlichen Pendants (vgl. Wagner 1984, S. 99).

Widersprüchlich bezüglich der typischen Kennzeichen der Rechts- und Verwaltungssprache sind die unterschiedlichen Ergebnisse, die empirische Untersuchungen zum Gegenstand produzieren. So hält Wagner (1984, S. 100) fest, dass kurze Sätze Merkmale von Verwaltungstexten sind, wohingegen Hansen-Schirra/Neumann (2004, S. 170) 30 Jahre später lange Sätze als kennzeichnend für die Verwaltungssprache bezeichnen. Auch bezüglich der Verwendung von Kunstwörtern herrscht Uneinigkeit:

³⁵ sprachlicher Handlungsträger

„Die Verwaltungssprache enthält nur wenige Kunstwörter; sie knüpft vielmehr an die Allgemeinsprache an, gibt aber dem Allgemeinwort häufig eine besondere fachsprachliche Bedeutung.“ (Wagner 1984, S. 99)

„... Komplexität der Darstellungsweise durch Derivationen (d. h. die Bildung neuer Wörter durch Anfügen von Endungen bzw. Vorsilben), Ketten einander untergeordneter Substantive, ...“ (Hansen-Schirra/Neumann 2004, S. 170)

Ob die Unterschiede an einer so starken Veränderung der Verwaltungssprache in dem Zeitraum oder einer zufälligen Heterogenität der untersuchten Textkorpora liegt, kann im Rahmen dieser Arbeit nicht beleuchtet werden, wäre jedoch sicher eine interessante Fragestellung.

Einigkeit herrscht wiederum darüber, dass sich in Rechts- und Verwaltungstexten viele formelhafte Wendungen und archaische Formen finden (vgl. Wagner 1984, S. 101; Hansen-Schirra/Neumann 2004, S. 170), welche für den ungeübten Leser das Verständnis erschweren können.

3.1.4. Leseprozesstheorien/Lesemodelle

Klassische Lesemodelle

Fraglich ist an dieser Stelle noch, wie sich die komplexen Prozesse und Situationen aus Kapitel 3.2.1 modellhaft abbilden lassen, so dass diese operationalisierbar werden und Aussagen über Kompetenzen erlauben.

Lesen als decodierende Kommunikationsmethode

Als erste Annäherung kann Schnotz' Modell (1994) – entwickelt in Anlehnung an das Organonmodell der Sprache von Bühler – dienen. Hier wird ein Text als eine Kommunikationsmethode betrachtet, welche zwei Teilprozesse der Kommunikation miteinander kombiniert: den Prozess des Meinens und den des Verstehens. Dieser Kommunikationszusammenhang ist in Abbildung 5 schematisch dargestellt. Ein Autor besitzt Wissen über einen Gegenstand, einen Ablauf oder eine Idee in Form einer mentalen Repräsentation. Dieses Wissen externalisiert der Autor durch das Verfassen eines Texts. Der Autor drückt dabei mit dem Text seine mentale Repräsentation des Gegenstands/Ablaufs/Idee und der jeweiligen Beschaffenheit aus. Ein Leser „versteht“ den Text, indem er anhand der gegebenen (Text-)Informationen eine mentale Repräsentation des Dargestellten bzw. entsprechendes Wissen konstruiert. Der Leser versucht also zu ermitteln, was der Autor gemeint hat (oder haben könnte). Je kongruenter die vom Leser konstruierte mentale Repräsentation und die ursprüngliche Repräsentation des Autors sind, als desto erfolgreicher kann die Textkommunikation zwischen Autor und Leser beschrieben werden.

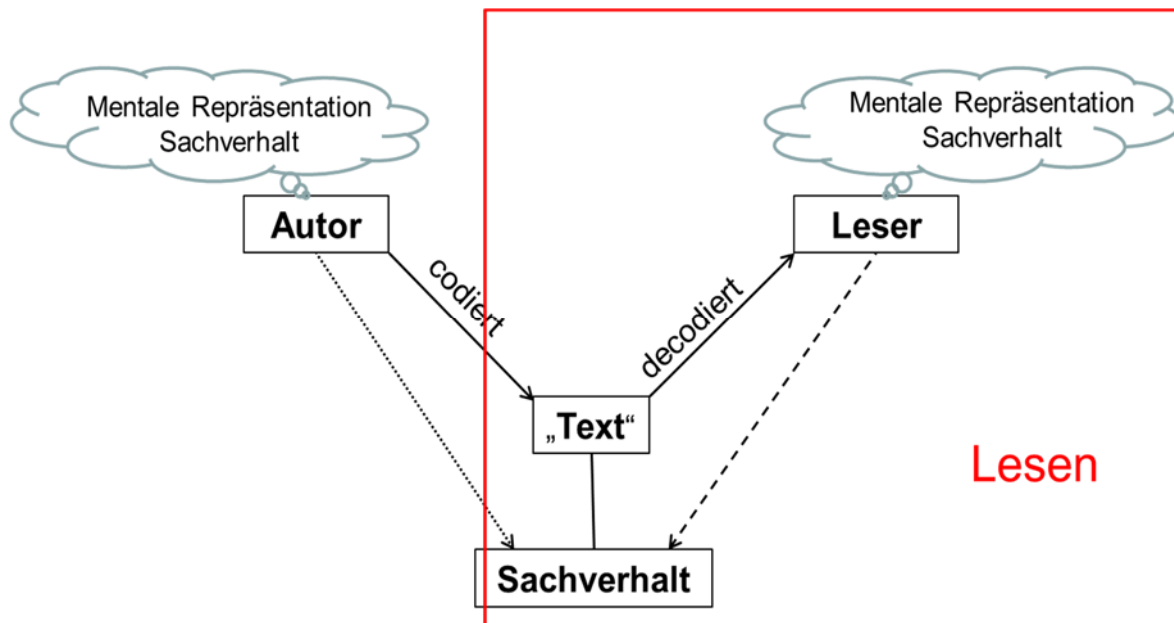


Abbildung 5 – Kommunikationszusammenhang Autor-Leser (Ziegler 2013 in Anlehnung an Schnotz 1994)

In Abbildung 5 ist der an dieser Stelle relevante Bereich – der des Lesens markiert. Im Weiteren wird davon ausgegangen, dass der Autor seine mentale Repräsentation adäquat abgebildet hat und somit eine Reduktion auf die Perspektive des Lesers gerechtfertigt ist.

In dem angesprochenen Modell werden die explizit ablaufenden Prozesse noch alle unter dem Stichwort decodiert subsumiert. Das ist hier nicht spezifisch genug, daher werden weitere Modelle zu Erklärung herangezogen.

Bereits bei der Sichtung gängiger Modelle fällt zunächst auf, dass unverbalisiert eine Einschränkung auf deskriptionale Repräsentationen vorgenommen wurde. Die Betrachtung von depiktionalen Repräsentationen im Kontext des Lesens fand bisher nicht die vermutlich angemessene Berücksichtigung.

Der Prozess des Lesens deskriptionaler Repräsentationen wird in der modernen Forschung als eine Text-Leser-Interaktion betrachtet. Lesen ist dabei einerseits ein textgeleiteter Konstruktionsprozess, andererseits ein vom Welt- und Sprachwissen gesteuerter, konzept- und erwartungsgeleiteter Integrationsprozess (vgl. Richter/Christmann 2002, S. 27).

Der Aspekt der Konstruktivität umfasst erstens, die sprachlich vermittelten Informationen einzelner Wörter und Sätze in einem emporsteigendem Verarbeitungsprozess mit den kontextuierten Informationen in Beziehung zu setzen (bottom-up-Prozess/-Verarbeitung) und zweitens, die erfassten semantischen und syntaktischen Strukturen in einem absteigenden Verarbeitungsprozess in eigene Erfahrungen, Vorwissensbestände und Emotionen zu integrieren (top-down-Prozesse/-Verarbeitung). Der Prozess des Lesens als textuelle Verstehensleistung ist somit durch eine kontinuierliche Reziprozität und Interaktion zwischen Textinformationen und Rezipient auf mehreren Ebenen gekennzeichnet (vgl. Groeben 1982, S. 48–50; Christmann/Groeben 2006, S. 147).

Detaillierte Erklärungsansätze zu den ablaufenden kognitiven Prozessen liegen zwar vor, allerdings existiert bis heute keine einheitliche kognitionspsychologische Theorie der Textverarbeitung. Es gibt diverse konvergierende, divergierende oder einander ergänzende Ansätze und Modelle, welche jeweils verschiedene Perspektiven des Verarbeitungsprozesses akzentuieren.

Die aktuelle Kognitionspsychologie ist zumindest d'accord in der Ansicht, dass Leser verschiedene Dimensionen oder Level der Repräsentation eines Textes konstruieren, so dass beim Textverstehen also von multiple mentale Repräsentationen ausgegangen werden kann (vgl. u.a. Kintsch/van Dijk 1983; Schnotz 1994; Zwaan 1992; Britton/Graesser 1996; Kintsch 1998). Unbestritten ist auch, dass Lesen ein komplexer und beziehungsreicher Vorgang der Bedeutungsaufnahme und -entnahme ist, welcher sich aus mehreren flexiblen Teilprozessen konstituiert, die sowohl auf der Wort- und Satz- wie auch der Textebene realisiert werden (vgl. Christmann/Groeben 2006).

Die unterste Ebene stellt die Dimension der Buchstaben- und Worterkennung dar und entspricht etwa einer mentalen Repräsentation der Textoberfläche. Dort können okulomotorische und perzeptuelle Prozesse verortet werden (d.h. sensorisches Erfassen von Buchstaben und Wörtern), gefolgt von Prozessen der Wortidentifikation, der phonologischen Dekodierung und der lexikalischen Zuordnung. Darauf aufbauend folgen semantische und syntaktische Verarbeitungen der Satz- und Textebene, die lokale Kohärenzen zwischen Wort-, Satz- und Textteilen etablieren (entspricht in etwa der propositionalen mentalen Repräsentation). Zur Bedeutungskonstruktion des Textes, der eigentlichen Sinnstiftung, wird dann von Prozessen ausgegangen, die lokale Kohärenzen mit globalen Kohärenzen verbinden und weiterführend Vorwissen und andere vorhandene Teilinformationen integrieren (vgl. Groeben/Hurrelmann 2002, S. 25). Wie diese Teilprozesse im Detail zu beschreiben sind und interagieren wird unten noch genauer beleuchtet.

Buchstaben- und Worterkennung

Das heute in der Kognitionsforschung etablierte und anerkannte Leseprozessmodell nach Walter Kintsch, wird, ungeachtet dessen, dass es nicht völlig unumstritten ist, aus theoretischen, wissenschaftskriterialen und forschungspraktischen Gründen auch hier zentral fokussiert, aber auch im Anschluss kurz hinterfragt.

Der Leseprozess entsteht dabei auf dem sogenannten hierarchieniedrigsten Level in einem primär visuellen Verarbeitungsvorgang, der der Buchstaben- und Wortidentifikation und Wortbedeutungserfassung dient. Die Transformation von Graphemen in Phoneme ist bedingt durch die Befähigung zum Dekodieren – was meint: die Fähigkeit des Entschlüsselns von Buchstaben und Wörtern muss entwickelt sein. Sie dient hauptsächlich der Unterscheidung von Analphabeten und alphabetisierten Individuen, da sie als pure Lesefertigkeit nicht genügt, um den Inhalt, die Bedeutung eines Satzes zu ermitteln (vgl. Richter/Christmann 2002, S. 29).

Umstritten ist in der Forschung, wie sich der Identifikationsprozess der Buchstaben und Wörtern gestaltet. Es werden zwei grundsätzlich verschiedene Paradigmen antizipiert:

Gough (1972) nimmt beispielsweise eine serielle Verarbeitungsweise an, bei welcher einzelne Buchstaben von links nach rechts nacheinander identifiziert werden. Smith (2004; erstmalig 1971) dagegen vermutet die Verarbeitung von Wörtern als ganzheitliche visuelle Muster. Keines der angenommenen Modelle konnte in seiner Alleingültigkeit eindeutig falsifiziert oder verifiziert werden. So konnte Cattell (1986) durch die Darstellung des Wortüberlegenheitseffekts bereits Hinweise zur Widerlegung der seriellen Verarbeitung von Buchstabenfolgen liefern. Aber auch Smith' Modell, der Speicherung von Wörtern als visuelle Muster, ist durch Untersuchungen, in denen Leser Wörter trotz typografischer Abweichungen von einem Standardformat identifizierten, zweifelhaft.

Somit wird derzeit, statt eines visuellen Identifizierungsprozesses auf Buchstaben- oder Wortebene, eine parallele Verarbeitung basierend auf abstrakten Buchstabeneinheiten angenommen (vgl. Coltheart 1981; Rayner/Pollatsek 1989).

Ein Beispiel für diese Betrachtungsweise ist das von McClelland/Rumelhart (1981) entwickelte interaktive Aktivationsmodell. Demzufolge kann der Wortidentifikationsprozess als Interaktion zwischen gegenseitiger Hemmung und Aktivierung von gespeicherten graphischen Eigenschaften, Buchstaben und Wörtern interpretiert werden.

Problematisch ist dieses Modell bei der Erklärung der Identifikation von Wörtern, die nicht als fertige Einheit in einem mentalen Lexikon eingetragen sind – dies wären bspw. Komposita, Flexionsformen oder unbekannte Wörter. Hierbei stößt das Modell offensichtlich an seine Grenzen, so dass eine Modifikation notwendig scheint (vgl. Christmann/Groeben 2006). Daher wird angenommen, dass es neben einem direkten Zugang für die im Lexikon gespeicherten Wörter auch einen indirekten Pfad vermutlich über das phonologische System gibt (vgl. Seidenberg et al. 1984).

Des Weiteren wird zur Erklärung auf das sequentielle Zugangsmodell verwiesen, welches auf der Kodierung morphologischer Strukturen von Wörtern gründet. Gemäß dieses Modells enthält das mentale Lexikon sowohl morphologische Formcharakteristika (auf der Grundlage abstrakter Buchstaben) und zusätzliche Inhaltsattribute. Der lexikalische Input wird dann im Prozess der Wortidentifikation in morphologische Einheiten zerlegt, bei dem zunächst das Stamm-Morphem im Lexikon identifiziert wird. Alsdann werden die Präfixe und Suffixe unter Rückgriff auf spezielle Regeln analysiert (vgl. Taft/Foster 1975; Taft 1986; Feldmann 1991). Diesbezüglich konnte bereits empirisch nachgewiesen werden, dass die Zeit, die für das Erkennen eines Wortes benötigt wird, abhängig ist von der Häufigkeit des Stamm-Morphems, nicht aber von der Häufigkeit des vollständigen Wortes (vgl. Taft 1979; Artelt et al. 2007, S. 18). Ebenso konnte die psychologische Relevanz der morphologischen Struktur durch Studien belegt werden, welche eine stärkere Voraktivierung („Stamm-Primingeffekt“ (Zwitserlood 1995, S. 116))

für morphologisch verwandte Wörter³⁶ als für morphologisch nicht-verwandte Wörter³⁷ aufzeigen konnte (vgl. Christmann/Groeben 2006, S. 150).

Zusammenfassend lässt sich konstatieren, dass sich dem Leser bei der Wortidentifikation drei sich nicht ausschließende Optionen bieten:

1. der direkte visuelle Zugang über eine Aktivationsausbreitung für Wörter, die bereits im mentalen Lexikon vorhanden sind
2. der indirekte Zugang über das phonologische System
3. ein Zugang über die morphologische Struktur der besonders für unbekannte und komplexe Wörter geeignet scheint (vgl. Christmann/Groeben 2006, S. 151).

Das bisher genannte bezog sich im Kern hauptsächlich auf die singuläre Wortebene. Auf höherer Ebene kann (immer noch bezüglich der Wortidentifikation) die Frage nach dem Kontext in dem die einzelnen Wörter eingebettet sind – einem möglichen Satzüberlegenheitseffekt – gestellt werden. Bezüglich dieser kam es wiederum zur Manifestation zweier kontroverser Theoriepositionen. Sie differieren primär in dem Punkt der Modellierung des Zusammenspiels der Ebenen des Leseprozesses – in Form der interaktiven vs. der modularen Ansätze (vgl. Christmann/Groeben 2006, S. 151).

Die interaktive Modelle wie u.a. das Logogen-Modell von Morton (1969) gehen in diesem Punkt von der Prämisse aus, dass sich die am Lesevorgang beteiligten Teilsysteme weitgehend parallel, synchron oder zumindest überlappend vollziehen und höhere Verarbeitungsprozesse bereits einsetzen, ehe die Verarbeitung auf der niedrigeren Ebenen vollständig abgeschlossen ist. Insofern wird unterstellt, dass der sprachliche Kontext die Wortidentifikation vereinfacht (vgl. Herrmann 1990, S. 297).

Vertreter der modularen Ansätze bspw. Fodor (1983) dagegen weisen die Parallelitätsannahme ab und postulieren prinzipiell autonom agierende Teilsysteme. Folglich setzen höhere Verarbeitungsprozesse erst dann ein, wenn die Prozesse der niedrigeren Ebene (wie der Wortidentifikation) vollständig abgeschlossen sind. Der lexikalische Zugriff ist dementsprechend gekennzeichnet durch das Merkmal einer informationellen Einkapseltheit in ein lexikalisches Modul, welches nicht mit anderen Modulen interagiert. Insofern sollte der Kontext gemäß der Modularitätshypothese nicht vereinfachend auf den lexikalischen Zugang wirken (vgl. Fodor 1983; Christmann/Groeben 2006, S. 151).

Diese theoretischen Annahmen bezüglich der konkreten Relevanz des Kontextes auf lexikalische Prozesse wurden verschiedentlich untersucht, wobei die empirischen Befunde uneinheitlich zu

³⁶ z.B. „Maler“ und „malen“

³⁷ z.B. „Maler“ und „tanzen“

sein scheinen. Simpson/Krueger (1991) konnten in Untersuchungen zu Homonymen³⁸ Hinweise darauf liefern, dass abhängig vom präsentierten Kontext die konvenierende Wortbedeutung schneller aktiviert wird als die nicht-konvenierende Bedeutung. Dies widerspricht auch der Modularitätshypothese, denn in dieser wird angenommen, dass alle denkbaren Bedeutungen zunächst aktiviert werden und so lange im Kurzzeitgedächtnis verbleiben, bis die Mehrdeutigkeit durch den Kontext aufgelöst werden kann (Desambiguierung) bzw. durch Verarbeitungsprozesse auf höherer Ebene austariert werden. Auch diese gegensätzliche Vermutung wird durch Studienergebnisse gestützt, die zeigen konnten, dass sich die Reaktionszeiten für das Identifizieren kontextangemessener oder -unangemessener Bedeutungen (für eine Desambiguierung) nicht unterscheiden und dass auch dann beide Bedeutungen aktiviert werden, wenn eine Bedeutung dominant ist (vgl. Rayner/Pollatsek 1989, S. 224–226).

Der Widerspruch lässt sich jedoch durch weitere Analysen auflösen. Diese können aufzeigen, dass nicht nur die Stärke und Art des Kontextes sondern auch die Häufigkeit der jeweiligen Wortbedeutungen Einfluss darauf nehmen, ob eine oder beide Wortbedeutungen aktiviert werden. Sollten der Kontext sowie die Häufigkeit der Wortverwendung bereits ausreichende Unterstützung für eine Interpretation liefern, wird nur eine der möglichen Bedeutungen aktiviert, sollten beide Faktoren zu unterschiedlichen Interpretationen führen, so werden zunächst beide Bedeutungen aktiviert. Dies zeigt, dass sich die kognitiven Prozesse bereits auf Wortebene als sehr viel flexibler darstellen, als zumeist von einzelnen Modellen angenommen wird (vgl. Christmann/Groeben 2006, S. 152).

Satzebene

Über die Wortebene hinaus gilt es die Satzebene zu betrachten. Oben wurde bereits auf die nicht abschließende Beurteilung des Kontexteinflusses verwiesen, jedoch bezogen auf die basale Ebene der Worterkennung. Im Folgenden geht es um die Prozesse der darauf aufbauenden Hierarchieebene der Satzverarbeitung.

Kognitionspsychologische Theorien stimmen heute darin überein, dass davon ausgegangen werden kann, dass Wortfolgen, angesichts ihrer semantischen Relationen, als aufeinander bezogen betrachtet werden können und somit wahrscheinlich in ein strukturiertes Gesamtgefüge integriert werden, indem vermutlich semantische und syntaktische Relationen unter Syntagmen hergestellt und zu Bedeutungseinheiten zusammengefügt werden. Der Bedeutungsgehalt von Texten wird nicht an der Oberfläche, sondern über die semantische Textbasis transportiert. Diese Basis kann beschrieben werden, als ein relationales Gefüge, dessen Grundelemente Prädikat-Argument-Strukturen – sogenannte Propositionen – sind. Eine Proposition kann als eine Art Elementarsatz verstanden werden, der aus Wortkonzepten zusammengesetzt ist, die an der Textoberfläche quasi Wörtern entsprechen (vgl. Bock 1978, S. 55).

³⁸ D.h. lexikalisch ambige Wörter wie *Kiefer* (Nadelbaum vs. Teil des Gesichtsschädels), *sieben* (Zahlwort vs. durch ein Sieb schütten), *Tau* (Niederschlag vs. Seil vs. griechischer Buchstabe).

Schiefele (1996) bezeichnet eine Proposition dementsprechend in Anlehnung an Kintsch als kleinste Bedeutungseinheit, welche als eigenständige Behauptung oder Aussage stehen kann und die semantische Struktur des Textes repräsentiert (vgl. Kintsch 1994; Schiefele 1996, S. 93)

Das Konzept der Prädikat-Argument-Struktur wurde erstmal im Rahmen der Kasusgrammatik als Teil der Generativen Semantik von Fillmore (1968) bzw. Chafe (1970) expliziert.

„Die semantische Struktur eines Satzes wird hier als tiefenstrukturelle Relation zwischen einem Prädikat (Zustände, Ereignisse, Eigenschaften) und den von ihm implizierten Argumenten (Objekte, Personen, Sachverhalte) beschrieben. Das Prädikat steht dabei im Zentrum des Satzes und legt die Anzahl der möglichen Argumente fest, die zu ihm in einer besonderen Kasusbeziehung stehen“ (zitiert nach Christmann/Groeben 2006, S. 153).

Die psychologische Bedeutsamkeit der Prädikat-Argument-Struktur liegt in ihrem Verständnis nicht als sprachliche, sondern als kognitive Bedeutungseinheit, welche die Struktur unseres Wissens von der Welt abbildet.

Die psychologische Relevanz von Propositionen konnte anhand verschiedener empirischer Untersuchungen verifiziert werden. So haben umfangreiche Analysen des Hierarchieeffekts (level effect) eine höhere Behaltensleistung für ranghohe Propositionen im Vergleich zu rangniedrigeren Propositionen nachweisen können (vgl. zusammenfassend u.a. Britton/Graesser 1996). Insgesamt kann die Befundlage zur psychologischen Relevanz von Propositionen und der hierarchischen Textstruktur überwiegend als konsistent und stabil angesehen werden. Sie unterstützt den propositionalen Ansatz insofern, dass allgemein Konsens darüber herrscht, dass Textverarbeitungsprozesse als hierarchisch-sequentielle Organisationsprozesse betrachtet werden (vgl. Christmann/Groeben 1996, S. 134)

Aus der Transformation eines Textes in eine propositionale Struktur folgt ebenfalls eine relative Nähe zur grammatisch-linguistischen Struktur des betreffenden Textes. Diese propositionale Repräsentation des Textes wird als sogenannte Tiefenstruktur von Sätzen bezeichnet, die sich beispielsweise darin äußert, dass Prädikate an der Satzoberfläche durch Verben, Adjektive oder Adverbien realisiert sein können (vgl. Christmann/Groeben 2006, S. 153).

Eine alleinige semantische Analyse reicht für den kognitiven Verstehensprozess immer noch nicht aus. Diese muss durch eine syntaktische Analyse erweitert und ergänzt werden. Daher hat die syntaxorientierte Psycholinguistik zur vollständigen Analyse von Sätzen eine Anzahl weiterer Segmentierungsvarianten – genannt Parser – entwickelt, um diese Strukturen weiter erforschen, beschreiben und verstehen zu können. Zu den anerkanntesten Parsern gehören beispielsweise die Übergangsnetzwerke von Woods (1970) und Wanner (1980) oder die minimale Anbindung und späte Schließung (vgl. Rayner/Pollatsek 1989, S. 245–247).

Auch hier manifestierten sich, ähnlich der Wortidentifikationstheorien, verschiedene Positionen über die Bedeutung der syntaktischen und semantischen Analyse im Rezeptionsprozess. Eine der Theorien – die autonome Syntaxtheorie von Garfield (1989) – folgt der Modularitätsthese

und geht davon aus, dass die syntaktischen Verarbeitungsprozesse den semantischen Prozessen vorausgehen und unabhängig von diesen erfolgen. Konträr dazu postuliert die interaktionistische Syntaxtheorie überwiegend synchron-parallele Verarbeitungsprozesse, deren Analyse der syntaktischen Strukturen durch den pragmatischen und semantischen Kontext sowie vom Weltwissen der Rezipienten geprägt wird (vgl. Christmann/Groeben 2006, S. 155).

Die moderne Psycholinguistik scheint sich also durch die Erkenntnisse neuerer Studien insgesamt auf eine Position zuzubewegen, die primär den Aufbau semantischer Sinnstrukturen betont, welche indes bei komplexen und uneindeutigen Sätzen durch eine syntaktische Analyse lanciert wird. Somit scheinen aktuell (2015) eher interaktionistische Thesen im Vordergrund zustehen.

Kohärenzbildung und Textebene

Psychologische Modelle des Textverstehens befassen sich insbesondere mit der Frage, wie Leser eine Gesamtrepräsentation des Textes aufbauen. Dabei gibt es zwei Hauptströmungen der Forschung: die eher textseitig orientierte Perspektive und die eher leserseitig orientierten Forschung.

Erstere versucht auf unterschiedlichen Analyseebenen verarbeitungsrelevante Textmerkmale und deren Einfluss auf das Lesen, Verstehen und Behalten von Texten zu identifizieren. Das Anliegen letzterer hingegen besteht darin, die im Rezeptionsprozess bedingten kognitiven Aktivitäten (Vorwissen, Erwartungen, Zielsetzungen) und Kompetenzen der Leser zu modellieren.

Da ein Text ein kohärentes Konstrukt darstellt und in Beziehung stehende Sachverhalte vermittelt, sind auch zwischen den Propositionen einzelner Sätze vergleichbare Zusammenhänge anzunehmen (vgl. Rickheit/Strohner 1993; Kintsch/van Dijk 1978). Somit ist die nächste zu betrachtenden Ebene die der lokalen Kohärenzbildung. Diese Kohärenzbildung geschieht durch mikropropositionale Verknüpfungen, die den semantischen Relationen zwischen den direkt folgenden Sätzen entsprechen. Der Prozess der Kohärenzbildung wird vereinfacht durch eine hohe Kohärenz des zugrundeliegenden Textes. Diese ist gekennzeichnet durch viele und eindeutige Hinweise, wie die Sätze des Textes sprachlicherseits durch lexikalische und grammatische Mittel miteinander verknüpft sind und im Leseprozess zu einer semantischen Einheit integriert werden sollen. Kohärenzbeziehungen können dabei via diverser semantischer und syntaktischer Methoden hergestellt werden. Eine der simpleren Formen von Kohärenz sind exemplarisch Rekurrenz (Wortwiederholung) oder die Wiederaufnahme von Satzteilen durch Deiktika – zusammengefasst unter dem Fachterminus Koreferenz (vgl. Christmann/Groeben 2006, S. 158). Die Auswirkungen von verschiedener Koreferenzarten auf kognitive Prozesse insbesondere unter der Berücksichtigung der Geschwindigkeit der Verarbeitung wurden in unterschiedlichen Studien untersucht und mehrfach als relevant aufgezeigt. Es konnte unter anderem gezeigt werden, dass sich besonders Wortwiederholungen im Folgesatz für die Verarbeitungsprozesse und die Geschwindigkeit dieser förderlich darstellen, sofern das jeweilige Wort im ersten Satz an der Objekt- und im zweiten Satz an der Subjektposition steht (vgl. Rickheit/Strohner 1993).

Noch globaler auf Strategieebene wird das Koreferieren und die syntaktische Integration bei der Betrachtung der Thema/Rhema-Progression (auch: Topic/Comment- oder Given/New-Strategie) u.a. von Hockett (1958) und Grimes (1984) erforscht. Dabei wird das Thema (Topic) eines Satzes als die schon geläufige, angeführte oder durch den Kontext transportierte Information bezeichnet und das Rhema (Comment) demgegenüber als eine Aussage über das etablierte Thema (Topic). Leichter zu erfassen ist dies an Hand zweier einfacher Beispiele:

1. In Darmstadt (Thema) gibt es die Mathildenhöhe (Rhema).
2. Es gibt Kuchen (Rhema). Dieser (Thema) ist aus Marzipan (Rhema).

Bei der Beschäftigung mit einem Thema konstruiert der Leser somit eine topicspezifische mentale Substruktur, die von Satz zu Satz (oder Teilsatz zu Teilsatz) beibehalten und sukzessive aggregiert wird, solange der betreffende Topic beibehalten wird (vgl. Bock 1978, S. 44–48; Christmann/Groeben 2006, S. 158f). Darüber hinaus können auch auf Ebene der globalen Textstruktur Signale zur Integration von Textinformationen transportiert werden. Dies kann zum einen durch so genannte Topic-Indikatoren, welche den Beginn eines Themas, seine Beendigung oder seine Fortsetzung indizieren, erfolgen, oder zum anderen mit Hilfe von sogenannten rhetorischen Relationen oder Signalen, mit denen die Stellung eines Themas im Gesamttext verdeutlicht werden kann (vgl. Meyer 1975). Artelt et al. führen dazu aus: „Im Prinzip geht es also darum, die unterschiedlichen Relationen zwischen verschiedenen Inhaltseinheiten (z.B. Ursache – Wirkung, Problem – Lösung, Grund – Folge, Frage – Antwort, These – Antithese explizit zu machen.“ (Artelt et al. 2007, S. 25). Empirisch bereits belegt werden konnte, dass das Einarbeiten von Hinweisen an der Textoberfläche zu einer qualitativ besseren Behaltensleistung und zu einer kohärenteren und vernetzteren Wissensstruktur führen kann (vgl. Artelt et al. 2007, S. 25).

Neben semantischen und syntaktischen Einheiten können auch mikro- und makrotypografische Darstellungsmethoden darauf verweisen, wie Textteile auf der Ebene der globalen Textorganisation aufeinander zu beziehen sind und zu globaler Kohärenz führen können (vgl. Ballstaedt et al. 1981; Groeben 1982; Günther 1988).

Zusammenfassend kann der Stand der Forschung zumindest als übereinstimmen darin bezeichnet werden, dass ausgehend von der bisherigen Befundlage Kohärenz auch als eine Eigenschaft des Textes aufgefasst wird, sondern als eine aktive leserseitige Konstruktion verstanden werden sollte (vgl. Schnotz 1994).

Die textorientierte Modellierung des Textverstehens fokussiert die Struktur eines Textes und versucht diese möglichst objektiv und präzise zu erfassen. Als Ausgangspunkt der modernen kognitionspsychologischen Modelle zu Textverarbeitungsprozessen wird häufig das Propositionsmodell von Kintsch (1974) bezeichnet, welches sowohl aus theoretischer, wie aus wissenschaftskriterialer und auch forschungspraktischer Perspektive besonders zu würdigen ist. Es wurde zunächst von Kintsch aus der Sprachwissenschaft entlehnt und anschließend zum Modell der zyklischen Verarbeitung von Kintsch/van Dijk (1978) sublimiert. Im ursprünglichen Modell

(Propositionsmodell) werden Texte in einem regelhaften Zergliederungsprozess durch die Bildung, Ordnung und reziproke Bezugnahme von Propositionen in eine Textbasis transferiert. So kann der jeweilige Bedeutungsgehalt einer Textrepräsentation expliziert werden und die Bedeutungsstruktur des Textes so objektiv wie möglich, was meint intersubjektiv annähernd präzise, abgebildet wird. Ausgehend von der Textbasis kann sodann vermutlich mit Hilfe semantischer Kohärenzrelationen eine hierarchische Textstruktur modelliert werden. Dabei kann die Verarbeitung der Propositionen flüssiger und leichter erfolgen, wenn eine höhere Kohärenz der Textbasis vorliegt. Alleiniges Kohärenzkriterium nach Kintsch/van Dijk (1978) ist hier die referenzielle Identität. Demnach können Propositionen als kohärent bezeichnet werden, sofern sie ein gemeinsames Argument haben (Argumentwiederholung) oder eine Proposition vollständig als Argument in eine andere eingebettet ist (Propositionseinbettung) (vgl. Christmann/Groeben 2006, S. 163).

Umstritten bzw. problematisch an diesem Ansatz ist unter anderem der oben ausgeführte Objektivitätsanspruch bei der Beschreibung der Textstruktur. Da subjektive Verstehensprozesse bei der Transformation eines Textes in Propositionen oder der Konstruktion der hierarchischen Textstruktur nicht völlig ausgeschlossen werden können, kann eine völlig objektive Beschreibung des Textes als verhältnismäßig schwierig angesehen werden (vgl. Christmann 1989, S. 114–116).

Vor allem aber widerspricht die Idee, den Bedeutungsgehalt eines Textes vergleichsweise präzise als Summe der Propositionen dieses Textes abbilden zu können, der Auffassung des Leseprozesses als aktiv-konstruktiver Prozess und Akt der Bedeutungsgenerierung – was besonders bei Ironie oder andern literarisch-rhetorischen Stilmitteln auffällt: Diese lassen sich propositional nicht abbilden.

Weiterhin problematisch gestaltet sich das von Kintsch/van Dijk gewählte Kohärenzkriterium der referenziellen Identität, gemäß dem Textkohärenz unbedingt nur auf Basis von Argumentaufnahme und -einbettung erzeugt wird. Dies belegen auch die Untersuchungen und Analysen von Black, Freeman und Johnson-Laird (1986). Sie legten Probanden verschiedene Texte mit identischer referenzieller Struktur vor und konnten feststellen, dass derjenige Text besser verarbeitet wurde, dessen Sätze im Hinblick auf das jeweilige Weltwissens wahrscheinlicher und glaubwürdiger erschienen. Damit ist hingegen nicht endgültig falsifiziert, dass die referenzielle Identität für die Etablierung einer kohärenten Textstruktur irrelevant ist – sie kann jedoch häufig als für die Textverarbeitung alleine unzureichend gehalten werden (vgl. Christmann/Groeben 2006, S. 164).

Auf Grund dieser Unzulänglichkeiten entwickelten Kintsch/van Dijk (1978, 1983) das zyklische Modell der Textverarbeitung, in dem Leseprozesse nicht mehr als ein unflexibler regeldominierter Vorgang beschrieben werden. Stattdessen wird Textverstehen als variabler, dynamischer und strategischer Prozess interpretiert, der auch extratextuale Aspekte wie Ziele, Erfahrungen und Vorwissensbestände des jeweiligen Lesers einkalkuliert. Dieser Prozess läuft dabei in separaten, mehrere Phasen umfassenden und sich teilweise parallel vollziehenden Zyklen ab. Im

ersten Zyklus wird abhängig von der lezerspezifischen Kapazität des Arbeitsgedächtnisses eine Einheit von 2-20 Textpropositionen – sogenannte Chunks – ins Arbeitsgedächtnis importiert und bezüglich ihrer Kohärenz geprüft. Ausgehend von den Propositionen dieses Chunks wird ein mentaler Kohärenzgraph erstellt. Um diesen zu konstruieren wird eine Top-Proposition gewählt, der die anderen Propositionen logisch-hierarchisch untergeordnet werden. Der sich so entwickelnde Teilgraph wird alsdann komplett ins Langzeitgedächtnis transferiert. Zudem werden weitere Propositionen selektiert, die parallel im Arbeitsgedächtnis aktiv gehalten und ins Kurzzeitgedächtnis (Kurzzeitbuffer) transferiert werden. Im darauffolgenden Zyklus wird versucht zwischen den neuen Propositionen des zweiten Chunks und den bereits im Kurzzeitgedächtnis vorhandenen Teilstrukturen Kohärenz herzustellen. Ist das erfolgreich, so fängt der Selektionsprozess erneut an. Am Ende des gesamten Prozesses sollte im Langzeitgedächtnis der vollständige Kohärenzgraph des Textes vorliegen. Ist es auf Grund von Verarbeitungsschwierigkeiten nicht der Fall, so müssen Kohärenzlücken durch zeitintensive und verarbeitungserschwerende Verarbeitungsprozesse wie Reinstatements³⁹, Umorganisationen⁴⁰ oder Inferenzen⁴¹ geschlossen werden (vgl. Christmann/Groeben 2006, S. 165).

Die Grenze dieses Modells ist weiterhin der Anspruch der Referenzidentität als alleiniges Kohärenzkriterium. Eine weitere Schwierigkeit liegt in der praktikablen Anwendbarkeit dieses Modells für Verstehensvorgänge längerer und komplexerer Texte. Die Repräsentation in einer propositionalen Mikrostruktur ist an dieser Stelle nicht mehr erklärungs mächtig genug, da einerseits die Gedächtniskapazität erheblich überlastet werden würde und andererseits das Verständnis eines umfangreicheren Textes nicht zwingend alle Detailinformationen erfordert.

Dies ist der Ausgangspunkt für die sogenannten Makrostrukturmodelle. Ihnen zu Grunde liegt die Überlegung, dass das Gelesene wegen der begrenzten kognitiven Verarbeitungskapazität jeweils nur partiell mental präsent ist. Demzufolge kann angenommen werden, dass Leser einen Text nicht Satz für Satz, Phrase für Phrase in Propositionen umformen, sondern basierend auf der propositionalen Textbasis synchron ebenfalls globalere semantische Einheiten (Makropropositionen) generieren. Diese sind im Gegensatz zu den detaillierteren elementaren Mikropropositionen der Textbasis vorangestellt und repräsentieren deren Bedeutung in komprimierter Form, so dass Leser bereits während des Lesens weniger Sätze eine holistische, integrierte Repräsentation des aus Mikro- und Makropropositionen entwickelten Textinhalts aufbauen. Durch den Prozess der Makrostrukturbildung kann das Arbeitsgedächtnis trotz seiner begrenzten Kapazität relativ umfassende mentale Teilstrukturen aufnehmen (vgl. Schnotz 1988, S. 301). Für

³⁹ erneute Aktivierung von Propositionen durch Absuchen der bereits ins Langzeitgedächtnis überführten Teilgraphen nach Propositionen, die konkrete/bessere Anknüpfungsmöglichkeiten bieten und wieder ins Kurzzeitgedächtnis geholt werden. Beim Lesen eines Textes geschieht dies häufig durch Rückwärtsbewegung (Sakkaden) im Text, also durch das erneute Lesen einer oder mehrerer Passagen.

⁴⁰ Vollständige oder partielle Restrukturierung des aktuell generierten Graphen, indem bspw. eine andere Top-Proposition als Organisationskern des Graphen gewählt wird.

⁴¹ Inferenzen können schlussfolgernd aus dem bisher gelesenen Text oder durch das aktivieren zusätzlichen Weltwissens gelöst werden. Sie sind insbesondere notwendig sofern der Text der zur Schließung der Kohärenzlücke erforderlichen Information nicht enthält.

die Bildung der Makropropositionen existieren vier angenommene Operationen für die semantische Informationsreduktion: Auslassen⁴², Selegieren⁴³, Generalisieren/Abstrahieren⁴⁴ oder Konstruieren⁴⁵. Neben den Makroregeln beinhaltet die Oberflächenstruktur eines Textes oftmals nützliche Hinweise bezüglich der Makrostrukturbildung. Zu ihnen gehören bspw. thematische Indikatoren wie Metaaussagen, Zusammenfassungskindikatoren, Relevanzindikatoren, Konnektoren oder Indikatoren des Themenwechsels (vgl. Christmann/Groeben 2006, S. 166).

An dem Konzept der Makrostruktur wird häufig die Kritik am gesamten Konzept angesetzt. Dabei geht es vor allem um die genaue Anwendung der Makroregeln. Diese sind zwar formal eindeutig formuliert, doch sind die zugrundeliegenden kognitiven Operationen noch relativ unscharf beschrieben. Fraglich bleibt, wie die Makroregeln konkret angewendet werden – bspw. scheint die Wahl relevanter bzw. der Ausschluss irrelevanter Propositionen eher intuitiv als regelgeleitet zu erfolgen und damit einer entscheidenden Subjektivität und Variabilität zu unterliegen (vgl. Kintsch/van Dijk 1978). Grund dafür könnte die bereits ausgeführte Text-Leser-Interaktion sein, da der Einsatz der Makroregeln nicht nur textbasiert (bottom-up) geschieht, sondern auch Beeinflussung von den individuellen Vorwissensstrukturen, Erlebnisse und Interessen (und Intentionen (vgl. Schmalhofer/Glavanov 1986)) des Lesers (top-down) erfährt. Diesbezüglich ist auch ein Problem, dass die Prozedur ein interpretatives Verfahren ist, da die Makroregeln nicht in einem Automatismus zu einem einzigen „korrekten“ Textthema führen und darüber hinaus auch etwas nicht im Text implizites abbilden können (vgl. Adamzik 2004, S. 130f; Brinker 2005, S. 52–54).

Leserfokussierende Modelle

Im Gegensatz zu den bisher angeführten Perspektiven konzentriert sich die leserseitig gerichtete Forschung auf die Aktivitäten der Rezipienten und welchen Einflüssen sie ausgesetzt sind beim Verstehen eines Textes. Etwa zeitgleich zu den vorgestellten propositionalen Modellen fokussieren die schematheoretischen Ansätze zur Textverarbeitung vornehmlich den Einfluss von (Vor-)Wissensstrukturen, Erwartungen und Intentionen der Leser bezüglich des Verstehens und Behaltens von Texten (vgl. Christmann 2000, S. 113; Adams/Collins 1977). Im Kern dieser Konzepte steht der Schema-Begriff, welcher in der Psychologie bereits eine längere Verwendung aufweist (vgl. Bobrow 1975, S. 132). Bobrow (1975) und Schank (1982) präzisieren das Schema-Konzept Ende der 70er, Anfang der 80er Jahre (Schank zusammen mit Abelson) als ein Konzeptkonglomerat, welches variierende Abstraktionsebenen bezüglich Wissen über bestimmte charakteristische Zusammenhänge von Lebenswelten (z.B. Bibliotheken, Computer, Filme, Autos etc.) aufweist (vgl. Bobrow 1975, S. 132f; Christmann 2000, S. 117).

⁴² Ignorieren von für das Verständnis des Textes vermutlich irrelevanter Propositionen.

⁴³ Selegieren von Propositionen sofern bekannte Bedingungen, Bestandteile, Präsuppositionen oder Folgen von Handlungen, Situationen etc. eingeschlossen und/oder impliziert sind.

⁴⁴ Auslassen von essenziellen, merkmalskonstituierenden Komponenten eines Konzepts bzw. Ersetzen von Informationen durch ein gemeinsames Superkonzept bzw. einen Oberbegriff.

⁴⁵ Bildung einer neuen Proposition, die die inhärenten Zusammenhänge zwischen den Konzepten der einzelnen Propositionen (Bedingungen, Folgen, Eigenschaften oder Komponenten einer Handlung u.Ä.) einschließt.

Über das Schema-Konzept lässt sich konsensual festhalten (wenn auch in der Kognitionspsychologie bezüglich einiger Details⁴⁶ uneinheitliche Positionen existieren (Bobrow 1975; Rumelhart 1980; Bahrick et al. 1988; Sadoski et al. 1991)) (vgl. Bobrow 1975, S. 132–136):

- Schemata setzen sich aus mentalen/kognitiven Wissensstrukturen, aus einer Konfiguration von Konzepten sowie deren Interrelationen, die jeweils Abstraktionen der ursprünglich dargebotenen Konzepte sind, zusammen.
- Schemata sind nach dem Abstraktionsgrad ihrer Begriffe assoziiert
- Es können Subschemata/-strukturen vorhanden sein.
- Es gibt Leerstellen (Slots), die durch eingehende Informationen ergänzt oder durch hypothetische Konzeptionen gefüllt werden können. So ermöglichen Schemata durch ihre Slots Eingliederungen neuer Informationen in die bereits vorhandene Wissensstruktur (Schemainstantiierung).
- Schemata führen zu Annahmen bezüglich der in dieses Schema integrierbaren Information und bestimmen so die Interpretation neuer Informationen.
- Schemata werden durch Folgerungen aus gemachten Einzelerfahrungen abgeleitet
- Sie sind prozessual durch eine Wechselbeziehung von datengeleiteten (bottom-up) und schemagesteuerten (top-down) Verarbeitungsaktivitäten charakterisiert.

Es lässt sich erkennen, dass Schemata im Textverarbeitungsprozess primär eine aufmerksamkeitssteuernde Funktion haben, die Integration und Interpretation neuer Informationen erleichtern und in der Abrufphase die Rekonstruktion gespeicherten Wissens steuern können (vgl. Adams/Collins 1977, S. 5–8).

Die Aussagekraft von Schemata für die Textverarbeitung wurde in einer Reihe von empirischen Analysen anhand von Texten mit ambigen Inhalten überprüft. Dabei wurden die Texte einmal mit einer sinnstiftenden Integrationshilfe (bspw. Texttitel und Thema) konzentriert und einmal ohne eine solche Hilfe offeriert. Die Ergebnisse konnten zeigen, dass bei der Vorgabe einer Integrationshilfe eine signifikant bessere Behaltensleistung erzielt wurde – über das Textverstehen wurden keine Aussagen getroffen. Dennoch kann angenommen werden, dass Integrationshilfen vorab Schemata aufrufen, die bei der Interpretation und Organisation des Textes helfen. Ähnliches zeigen auch Untersuchungen mit vorgegebenen Leseperspektiven. Vermutlich beeinflussen diese die Perzeption und Impression relevanter Textelemente. Abhängig von der indu-

⁴⁶ Wie der Unterschiede zwischen Schemata, Konzepten, Skripte und Geschichtengrammatiken, daher wird im Folgenden, zur Darstellung der Theorie nur der Begriff Schemata verwendet, wenngleich damit eine Unschärfe bestehen bleibt.

zierten Leseperspektive werden unterschiedliche Schemata aktiviert und somit andere Textinformationen als relevant wahrgenommen und später auch erinnert (vgl. Schwarz/Flammer 1979).

Die schemaorientierte Theorie bietet in der modernen Linguistik und Pragmatik einen integrativen Rahmen für verschiedene Konzepte bezüglich der Beschreibung mannigfaltiger Arten von Wissensstrukturen. Dazu gehören exemplarisch Skripts und Geschichtengrammatiken, an dieser Stelle für das Textverstehen als zwei Spezifikationen des Schemabegriffs gelten können.

Während Schemata als auf die Abbildung allgemeinen Weltwissens fokussiert verstanden werden können, verkörpern Skripts und Geschichtengrammatiken spezifischeres Wissen. Schank und Abelson (1977) haben das Konstrukt der Skripts als routinierte Ereignis- und stereotype Verhaltensabläufe beschrieben. Skripts wie bspw. ein Restaurantbesuch dabei sind hierarchisch organisiert und weisen wie Schemata Slots auf, welche im Verarbeitungsprozess durch musterhafte Referenzen (Objekte und Rollen) – im Falle des Restaurants bspw. Tisch, Speisekarte, Essen etc. und Gast, Kellner, Koch etc. – gefüllt werden. Unter Rekurs auf entsprechendes Skriptwissen ist es einem Leser somit möglich, trotz mangelnder Informationen einen Text zu rekonstruieren bzw. zu verstehen. Den empirischen Beleg dafür lieferten Abbott, Black und Smith 1985 (vgl. Christmann/Groebe 2006, S. 168).

Des Weiteren unter dem Begriff Schematheorie subsummiert wird die sogenannte Geschichtengrammatik („story grammars“), welche die globale Systematik und Struktur von Textelementen bei literarischen Prosaexten beschreibt. Die diesbezüglich entwickelten Grammatiken konstituieren aus Regeln, die die Eckpfeiler einer Erzählung wie Thema, Setting, Ereignis, Charaktere etc. und diesbezügliche hierarchische und sequentielle Positionen in der Gesamtstruktur spezifizieren. Es kann postuliert werden, dass Geschichtengrammatiken im Kognitionssystem schemafundiert sind, d.h. Erzähltexte werden unabhängig von den konkreten Inhalten so interpretiert und erinnert, wie es in den entsprechenden Grammatiken spezifiziert wird. Diese Annahme konnte an Hand verschiedener Studien verifiziert werden, welche die Darbietungsfolge eines nicht optimalen Erzähltextes mit seiner Reproduktionsfolge verglichen haben. Dabei wurde gezeigt, dass bei mehrfacher Präsentation der (re)produzierte Text der Probanden immer stärker vom Ursprungstext differierte⁴⁷. Es konnten dabei auch zwei Hypothesen bezüglich der kognitiven Organisation des Gelesenen – die „Hierarchiehypothese“ sowie die „Ordnungshypothese“ – empirisch nachgewiesen werden, jedoch zeigte sich ebenfalls, dass das Verstehen und Behalten nicht einfach nur als eine Funktion der hierarchischen und sequentiellen Struktur der Textelemente modelliert werden kann (vgl. Christmann/Groebe 2006, S. 168).

Es lässt sich zusammenfassen, dass in den schematheoretischen Ansätzen die Relevanz des Vorwissens erstmalig deutlich als eine zentrale Einflussvariable im Rezeptionsprozess eindrucksvoll herausgearbeitet werden konnte. Jedoch wurde parallel auch die Problematik offenkundig, die der Versuch die Vorwissensstrukturen von Rezipienten zu modellieren mit sich bringt – insbesondere, wenn das Ziel ist, zuverlässige Vorhersagen darüber zu treffen, wie Texte verstanden

⁴⁷ Vgl. das Kinderspiel „Stille Post“

und behalten werden können. Dies ist der Hauptgrund, weshalb die Schematheorie stark umstritten ist. Ein weiterer zentraler Kritikpunkt ist die mangelnde Flexibilität und Kontexthypersensibilität der Modelle, die notwendig für die Beschreibung der Nutzung von Wissensbeständen im Prozess der Textverarbeitung ist (vgl. Waldmann 1990).

Synthese: Mentale Modelle

Als Synthese beider Ansätze kann die Theorie der mentalen Modelle bezeichnet werden, die sowohl die Wechselwirkungen zwischen propositionalen Strukturmomenten des Textes, lokalen und globalen Kohärenzen sowie dem Vor- und Weltwissen des Rezipienten zu integrieren versucht.

Diese wurde aufbauend auf den Forschungen der 70er und 80er Jahre entwickelt und versuchen damit einen integrativen Rahmen für die umfassende Modellierung der Interaktion von textbasierten Informationen mit Weltwissen und nicht-sprachlichen Vorwissensstrukturen zu bilden. Die Theorie weist im Laufe ihrer Entwicklung eine große Bandbreite an theoretischen Ansätzen und vor allem auch heterogene Begrifflichkeiten (referenzielles Modell, Situationsmodell, discourse model oder Szenariorepräsentation – welche in der Literatur jedoch weitgehend synonym verwendet werden) auf (vgl. bspw. Johnson-Laird 1983; Kieras/Bovair 1984; Glenberg et al. 1987; Schnotz 1988; Zwaan 1992; Schnotz 1994; Garnham/Oakhill 1996; Kintsch 1998, 2004). Auch dieses Modell stellt kein geschlossenes Konstrukt der modernen Textverstehensforschung dar. Dennoch besteht jedenfalls allgemeiner Konsens bezüglich des Verständnisses mentaler Modelle als analoge, inhaltspezifische und anschauliche Repräsentation(en) des im Text beschriebenen Sujets, deren Qualität auch über das erreichte Textverständnis Auskunft geben kann (vgl. Richter/Christmann 2002, S. 34).

Schnotz (1994) führt dazu aus:

„Wie bei anderen Modellen beinhaltet die Analogierelation auch hier, dass bestimmte Merkmale des mentalen Modells bestimmten Eigenschaften des Originals entsprechen. Sie beziehen sich auch hier nicht auf sämtliche Merkmale von Modell und Original. ... Die Eigenschaften des mentalen Modells müssen denen des Originals lediglich funktional analog sein. Es bedarf also keiner physikalischen, sondern lediglich einer natürlichen Isomorphie zwischen Modell und Original“ (vgl. Schnotz 1994, S. 158f.).

In der Theorie Mentaler Modelle wird also davon ausgegangen, dass diese den zu repräsentierenden Sachverhalt zwar auf der Grundlage der Textinformation als funktionale und strukturelle Analogie konstruieren, diese Konstruktion dann aber in einen realitätsabbildenden Sachverhalt transformiert wird. Während eine propositionale Repräsentation bei einer relativen Nähe zur linguistischen Struktur des Textes verbleibt, repräsentieren mentale Modelle den darzustellenden Sachverhalt direkt auf der Grundlage eines analogen Abbildungsprinzips: „Mental Models represent what the text is about, not the text itself“ (Glenberg et al. 1987, S. 70)⁴⁸.

⁴⁸ „Mentale Modelle bilden die Bedeutung des Textes ab, nicht den Text selbst“

Damit ist nicht gemeint, dass mentale Modelle den jeweiligen Sachverhalt in jedem Fall gänzlich repräsentieren. Das mentale Modell kann abhängig von Vorwissensbeständen, Intentionen und Interessen der Rezipienten den entsprechenden Realitätsausschnitt entweder in mehr oder weniger elaborierter oder auch in aspekthaft verkürzter Weise abbilden. In der Regel kann davon ausgegangen werden, dass das spezifische Segment fragmentarisch oder simpler ist als der reale Sachverhalt, der repräsentiert wird (vgl. Johnson-Laird 1983, S. 10).

Theoretisch erfolgt die Konstruktion mentaler Modelle sowohl über die Repräsentation der Texte auf der Ebene der propositionalen Abbildung als auch direkt auf der Ebene der mentalen Modelle. Es wird davon ausgegangen, dass sich diese beiden Ebenen im Verarbeitungsprozess ergänzen. Das mentale Modell wird einerseits durch die propositionale Struktur aktiviert und im Verarbeitungsprozess unter Rückgriff auf Vorwissensstrukturen allmählich elaboriert und ausdifferenziert. Die propositionale Struktur kann dabei als instruierend für die Modellkonstruktion verstanden werden, determiniert diese jedoch nicht abschließend (vgl. Christmann/Groeben 2006, S. 170).

Mentale Modelle gehen über strukturelle, topologische Gegebenheiten hinaus und eröffnen auch die Möglichkeit dynamische Sachverhalte zu repräsentieren. Sie befähigen einerseits zur Inferenzbildung und bieten andererseits auch die Möglichkeit Vorhersagen zu treffen, Phänomene zu verstehen, sich für eine Alternative zu entschließen und Handlungsausführung zu planen sowie zu kontrollieren. Vor allem ermöglichen sie es Ereignisse stellvertretend zu erfahren, da sie es erlauben, durch sprachliche Expressionen Repräsentationen zu erzeugen, die mit denen zu vergleichen sind, die durch eine direkte Kenntnis von Welt gewonnen werden kann (vgl. Christmann 1989, S. 87). Mentale Modelle können somit in gewisser Weise als figurativer, gegenständlicher und vorstellbarer angenommen werden als Propositionale Repräsentationen. Damit bieten sie darüber hinaus die Möglichkeit zu einem kognitiv recht komplexen Prozess: der mentalen Simulation und dem Imaginieren von Taten und Folgen – das gedankliche Probehandeln.

Als eines der wahrscheinlich elaborierteste Modell dieses Verständnisses des Textverarbeitungsprozesses kann das Modell von Kintsch/van Dijk (1983) gelten – das Situationsmodell. Damit wurde ein Konzept erarbeitet, welches den Prozess der Textrezeption als Wechselwirkung zwischen einem vorgegebenen Text und der Kognitionsstruktur der Leser versteht. Demgemäß wird der Text zwar entsprechend der zuvor deduzierten Propositionsanalyse repräsentiert, aber darüber hinaus wird ein Situationsmodell konstruiert, welches in gleichem Maße textbezogene Informationen und leserseitige Wissensbestände enthält. In diesem Modell wird Textverstehen als ein dynamischer, konstruktiver und interpretativer Prozess skizziert in welchem Leser in einem reziproken und sukzessiven Prozess („online-Annahme“) die Bedeutung eines Textes unter Rekurs auf sachbezogenes Weltwissen, individuelle Erwartungen, Erfahrungen und Intentionen (präsuppositive Annahme) konstruieren (vgl. Schnotz 1988, S. 305; Christmann/Groeben 2006, S. 170).

In diesem Kontext relevant ist noch die Modellierung des Textverarbeitungsprozesses als strategischer Prozess. Das wurde im Strategiemodell von Kintsch/van Dijk (1983) besonders berücksichtigt und weiter ausdifferenziert. Die Konstruktion des mentalen Modells erfolgt hierbei nicht willkürlich, sondern strategisch. Mit Strategien sind hier keine starren regelgeleiteten Prozesse gemeint, sondern stattdessen vielmehr Arbeitshypothesen, welche auf der Grundlage einer flexiblen Nutzung interner und externer Wissensbestände ausgebildet werden.

Ein weiterer Vorteil zu den oben genannten (Figuration, Gegenständlichkeit, Vorstellbarkeit) bezüglich der Annahme mentaler Modelle ist, dass die Konstruktion eines mentalen Modells insbesondere bei wissenschaftlichen Texten durch die Verwendung von Analogien und Metaphern gestützt werden kann, woraus sich positive Effekte für das kurz- und langfristige Behalten, auf Problemlösungen und die Güte der Beantwortung schlussfolgernder Fragen ergeben haben. Dies konnte auch bereits empirisch von Halpern/Hansen/Riefer (1990) nachgewiesen werden (vgl. Christmann/Groeben 2006, S. 170).

Trotzdem ist ein empirischer Nachweis analoger mentaler Modelle höchst schwierig, da die Analyse des Informationsgehaltes einer mentalen Repräsentation keinen Rückschluss darauf ermöglicht, ob es sich dabei um ein mentales Modell handelt oder nicht (vgl. Schnotz 1994, S. 194).

Bedacht werden muss jedoch auch, dass der Mentale-Modell-Ansatz in der Textverstehensforschung – insbesondere in den 90er Jahren – diskutiert und bezweifelt wurde. Einige Forscher sind der Ansicht, dass das Resultat einer Textverarbeitung – sofern der Leser keine speziellen Verarbeitungsstrategien verwendet – maßgeblich eine Repräsentation der propositionalen Struktur des Textes ist. Nur an bestimmten Positionen werden unter Rückgriff auf Weltwissen weitere Informationen eingefügt, sofern es sich um solche Informationen handelt, die für die Herstellung lokale Kohärenz erforderlich sind oder die besonders naheliegend und geläufig sind. Dieser Theorieansatz wird meistens als „minimalistische Hypothese“ (vgl. McKoon/Ratcliff 1992) bezeichnet (vgl. Kelter 1995, S. 71). Dies stellt sodann die Funktion eines Mentalen Modells für das Textverstehen grundsätzlich in Frage.

Gemäß dieses Theorieansatzes ist das Konstrukt des mentalen Modells für die Textverstehensforschung sodann größtenteils überflüssig. Dementsprechend sei ein Großteil der Forschung der Mentalen-Modell-Theorie darauf ausgerichtet die Notwendigkeit des Konstrukts für die Textverstehensforschung nachzuweisen. Somit sei dann das eigentliche und vorrangige Ziel der Forschung zu beweisen, dass es Textverstehensphänomene gebe, welche mit der Annahme von Repräsentationen propositionaler Strukturen nicht hinreichend aufgeklärt und interpretiert werden können und so das Postulat mentaler Modelle notwendig mache (vgl. Kelter 1995, S. 71f). Um diesen Nachweis führen zu können, wurden spezielle Methoden zur empirischen Differenzierung zwischen Repräsentationen propositionaler Textstrukturen und mentalen Modellen entwickelt. Nach der These von Johnson-Laird (1983), dass ein mentales Modell eine strukturell analoge Repräsentation des geschilderten Sachverhalts sei (vgl. Johnson-Laird 1983, S.

154f; 419f) lässt sich für bestimmte Texte prüfen, ob lediglich eine propositionale Repräsentationsstruktur oder zusätzlich bzw. stattdessen mentale Modelle gebildet werden. Wird bspw. festgestellt, dass zwei im Text genannte Entitäten⁴⁹, die in dem explizierten Sachverhalt räumlich in der Nähe liegen, mental enger miteinander verknüpft sind als zwei Entitäten, die in dem Sachverhalt in großer Distanz zueinander liegen, obwohl die „Abstände“ zwischen den Entitäten in der propositionalen Textstruktur jeweils dieselben sind, so ließe sich dies als Beleg für die Konstruktion eines mentalen Modells werten (vgl. Kelter 1995, S. 71).

Die diesbezüglichen empirischen Untersuchungen erbrachten verhältnismäßig klare Evidenz für Repräsentationen, die strukturell dem beschriebenen Sachverhalt analog sind. Vertreter der minimalistischen Hypothese bezweifeln jedoch auch nicht, dass zu Texten überhaupt mentale Modelle gebildet werden können. Vielmehr gemeint ist, dass zwar theoretisch mentale Modelle zu Texten gebildet werden können und dies auch des Öfteren vorkomme, es bedeute allerdings nicht automatisch, dass diese auch für gewöhnlich bei der Textrezeption konstruiert werden oder, dass mentale Modelle in summa eine wesentliche Funktion für das Textverstehen aufweisen. Autoren, welche der minimalistischen Hypothese zugeneigt sind, folgen der Auffassung, dass die in den bisherigen Untersuchungen nachgewiesenen mentalen Modelle vermutlich das Produkt spezieller, durch den jeweiligen Versuch, induzierter Verarbeitungsstrategien oder Weiterverarbeitungsprozesse gewesen seien, welche mit normaler Textverarbeitung nur minimal korrespondieren. Dieser Einwand scheint plausibel, wenn berücksichtigt wird, dass bei vielen Untersuchungen die Texte und/oder die Prüfmethode inhaltlich so strukturiert waren, dass die Versuchspersonen identifizieren konnten, welches die aufgabenrelevanten Textinformationen waren (bspw. die Informationen über räumliche Relationen) und somit gesonderte Verarbeitungsstrategien benutzen konnten, um diese Information korrekt zu erfassen und mental gut verfügbar zu halten (vgl. Kelter 1995, S. 71f).

Im Fokus alternativer Ansätze solle daher nach Kelter statt der Untersuchung, ob überhaupt mentale Modelle gebildet werden, eher die Frage stehen, ob die Bildung mentaler Modelle ein integraler Bestandteil des Textverstehens ist oder ob mentale Modelle nur unter Rückgriff auf besondere Verarbeitungsstrategien konstruiert werden. In den theoretischen Diskussionen wird dazu zunächst erörtert, ob und wie die für das Textverstehen obligatorischen Verarbeitungsprozesse (z.B. Erkennen der Wörter) von fakultativen Textverarbeitungsprozessen (z.B. wie hier angenommen mentale Modellbildung) separiert werden können (vgl. Kelter 1995, S. 71f).

Die Konfrontation von automatischer und kontrollierter Verarbeitung, wie sie in einigen Forschungsrichtungen üblich ist, scheint, so Kelter (1995), hier unbrauchbar zu sein, denn es kann nach dem bisher dargestellten nicht sinnvoll angenommen werden, dass Texte zumeist oder immer automatisch verstanden werden. Die Hypothese des automatischen Verstehens wird in dieser Form zwar auch bei der Annahme mentaler Modelle nicht gemacht, jedoch postulieren Vertreter dieses Ansatzes, dass die Bildung mentaler Modelle ein unabdingbarer Bestandteil des

⁴⁹ Hier: Informationsobjekte

Textverstehens sei. Sollten also Ergebnisse vorliegen, die zeigen, dass mentale Modelle nur unter speziellen Umständen konstruiert werden, unter anderen allerdings nicht, so spräche dies noch nicht gegen die gesamte Theorie Mentaler Modelle. Diese Resultate wären erst dann als Gegenbeleg im Sinne einer Falsifikation der Theorie zu werten, wenn zugleich nachgewiesen werden könnte, dass die Rezipienten ohne die Bildung mentaler Modelle – also durch die Verarbeitung auf einer rein propositionalen Ebene – den Text in unterschiedlicher Qualität verstanden hätten (vgl. Kelter 1995, S. 72).

Problematisch an diesem Konstrukt ist dann folglich die Konkretisation dessen, was „Verstehen“ umfassen soll. Ein Text kann vom Rezipienten mehr oder minder umfassend und richtig verstanden werden, genauso wie das erworbene Verständnis des Textes subjektiv für den Leser mehr oder minder sinnvoll und aussagekräftig sein kann. Die Frage nach dem Begriff des Textverstehens und der entsprechenden theoretischen Präzisierung sowie nach den Kriterien, welche in empirischen Untersuchungen bestimmen, ob ein Rezipient einen Text (oder eine Textinformation) verstanden hat oder nicht, sind für die Betrachtung jeglicher Ansätze zur Erklärung des Leseprozesses von Bedeutung (vgl. Kelter 1995, S. 72f).

Abschließend lässt sich festhalten, dass der aktuelle Forschungsstand weder unumstößliche empirische Belege für die Bildung mentaler Modelle liefert, auf der anderen Seite jedoch auch keinerlei endgültige dagegen hervorgebracht hat – es geht also im Kern nicht mehr um ein „Ob“, sondern, ob sie obligatorisch oder fakultativ gebildet werden. So lässt sich zumindest sagen, dass die Annahme einer strategiegeleiteten mentalen Modellkonstruktion plausibel erscheint (vgl. Schnotz 1994) und durch verschiedene empirische Arbeiten untermauert wird. Insofern muss die Theorie der mentalen Modelle zwar weiterhin ein hypothetisches Konstrukt bleiben, dennoch ist sie die aktuell schlüssigste Methode, welche die Interaktion mehrerer angenommener Komponenten des Textverstehens unter Berücksichtigung der kognitiven Flexibilität der Prozesse der Rezipienten annähernd adäquat modelliert (vgl. Christmann/Groeben 2006, S. 172). Ihre weiteren Vorteile zeigen sich im Verlauf der Arbeit bei der Betrachtung der Prozesse der Bildverarbeitung und Speicherung und der Integration beider Theorien und besonders bei ihrer Voraussetzung für das weitere Handeln aufbauend auf textbasierten Informationen.

Bild- und Grafikverarbeitung

Kommunikation mit depiktionalen Strukturen

Sprache und Bild sind in der westlichen Gesellschaft auf mannigfaltige Weise miteinander vernetzt und die Etablierung und Evolution der Bilder und die Ausbildung neuer Bildformate hat historisch gesehen das System des sprachlichen Handelns in wichtigen Teilbereichen modifiziert und differenziert (vgl. Muckenhaupt 1986, S. 1).

Straßner geht sogar so weit zu behaupten:

„Die mediale Kommunikation wird heute, im sogenannten ‚Optischen Zeitalter‘, nach der ‚visuellen Zeitenwende‘ bzw. im Zeichen der ‚Bilderflut‘, beherrscht von den Text-Bild- bzw. Bild-Text-Medien. Sie haben sich seit der Wende vom 19. zum 20. Jahrhundert überall durchgesetzt

(,pictorial turn') und meist eine Leitfunktion übernommen. In ihnen wird zum Teil versucht, neben dem Text eine ,Bildsprache' einzusetzen, die als ,visual esperanto' oder als Anhäufungen von ,Visiotypen', d.h. Stereotypen auf der Bildebene (vgl. Pörksen 1997) Allgemeinverständlichkeit zu garantieren scheint.“ (Straßner 2002, S. 1)

Es wird zuweilen Skepsis oder Unbehagen gegenüber dem massenhaften Gebrauch des medialisierten Bildes in der alltäglichen Kommunikation laut. Diese kritische Haltung kann mit dem in wissenschaftlichen Kontexten zu findenden Logozentrismus – der Überlegenheit von Sprache und Schrift gegenüber dem Bild – in Verbindung gebracht werden. Bilder wurden und werden häufig als oberflächlicher Blick auf den Anschein und als emotionalisiert betrachtet und mit mythisch-rituellen Handlungen verknüpft (vgl. Stöckl 2004, S. 7).

Die Sprache gilt vermutlich zu Recht als das zentralste und intensivste Mittel menschlicher Kommunikation. Offenkundig ist jedoch, dass sie nicht das einzige Mittel gegenseitiger Verständigung ist. Wenn jedoch Bedeutungen sprachlich nicht adäquat transportiert werden können oder wenn der Eindruck entsteht, dass diese nicht wie gewünscht verstanden werden, so besteht auch die häufig verwendete Möglichkeit, diese Botschaften durch Bilder oder Bildfolgen anschaulich zu machen – zu visualisieren. Das bedeutet für die Gesellschaft insgesamt den Übergang von einer weitestgehend schriftorientierten Kultur zu einer Kultur der Bildpräsenz und der visuellen Diskurse (vgl. Muckenaupt 1986, S. 156; Straßner 2002, S. 1).

Auf Grund der Durchdringung der Gesellschaft mit visueller Kommunikation entsteht zunehmend die Forderung nach einer Wissenschaftsdisziplin der „Imaging Science“ oder „Visualistik“. Insbesondere kommt der Wunsch nach einer „Visual Education“ oder „Visual Literacy“ als pädagogisch-didaktische Disziplinen auf (vgl. Wiese 1978; Straßner 2002, S. 2). Warum das notwendig sein könnte, wird deutlich, wenn die Perspektive auf logische Bilder fällt. Sie sind nach Meinung einiger Forscher kognitiv unentbehrlich, da sie kommunikative Aufgaben effizienter lösen, als es die Sprache zu tun vermag (vgl. Stöckl 2004, S. 7).

Wenn Bilder als Mittel der Kommunikation betrachtet werden, ist auffällig, dass es sich dabei um Instrumente handelt, die nicht wie die menschliche Sprache quasi natürlich und selbst existent sind, sondern um Methoden, die je nach Bedarf erst artifiziell geschaffen werden müssen. Die Disponibilität von und über Bilder ist daher weit limitierter als die Disponibilität über die Sprache. Häufig werden Bilder inkorrekt autonom von dem jeweiligen Produktionszusammenhang betrachtet, obwohl grade in der Reflexion der Produktionsprozesse, die zu den Bildern führen, ein wesentlicher Faktor für das Verstehen dessen, was Bilder sind, liegt. Eine grundlegende Unterscheidung für das Verständnis von kommunikativen Handlungen mit Bildern liegt darin, dass einerseits in der Produktion eines Bildes eine kommunikative Handlung mit Bildern vollzogen werden kann oder dadurch, dass bereits existierende Bilder in neuen Handlungskontexten integriert werden. Hinter dieser Formulierung – Bilder in Handlungskontexte integrieren – steht in vielen Fällen eine umfangreiche Auseinandersetzung mit den jeweiligen Bildern, die annäherungsweise so komplex ist wie das Produzieren der Bilder selbst. Beispiele dafür können

Filmschnitte und das Arrangement von Bildern in Werbeanzeigen sein (vgl. Muckenhaupt 1986, S. 156f), aber auch die didaktische Aufbereitung von Bildern für Lehr-Lern-Kontexte.

Somit wird deutlich, dass das Verstehen und der Umgang mit Bildern, als Mittel der Kommunikation – insbesondere unter der Perspektive des kognitiven Verarbeitungsprozesses – relevant für die Betrachtung ist.

Forschungsansätze und -richtungen

Zunächst fällt auf, dass die kognitive Verarbeitung von Bildern nicht in dem Ausmaß erforscht ist wie die Prozesse der Textverarbeitung. Während das Textverstehen in der kognitiven Psychologie bereits ausgiebig erforscht wurde, war das Bildverstehen ein eher randständiges Thema. Diese Tatsache rührte aus der fälschlichen Annahme, dass sich die Prozesse bei der Bildwahrnehmung nicht wesentlich von denen der visuellen Wahrnehmung unterscheiden. Aus der Neuropsychologie war bekannt, dass das natürliche Sehen und die Bildbetrachtung in denselben neuronalen Arealen stattfinden und es keine eigenen Module für die Bildverarbeitung gibt. Unterdessen läuft die Verarbeitung der gesprochenen Sprache und der Schrift in gesonderten Modulen ab. Somit kam man zu der Vermutung, dass das Bildverstehen nur ein Sonderfall des Sehens mit reduzierten visuellen Vorgaben sei. Demgegenüber wurde davon ausgegangen, dass viele Erkenntnisse aus der Textverstehensforschung bspw. über die Wirkung von Schemata oder über kohärenzbildende Prozesse einfach transferierbar seien. Unter Anwenderperspektive werden Bilder aktuell noch immer als leicht oder sogar selbstverständlich gegenüber der „komplexen“ Sprache verstanden. In der technischen Kommunikation wird eine derartige Zuversicht in die bildliche Kommunikation gesetzt, dass sprachfreie oder zumindest spracharme Anleitungen propagiert werden. Dementsprechend herrscht eine Ungleichgewichtigkeit beider Vermittlungsformen: während Textproduktion und -rezeption bereits in der Grundschule intensiv geübt und gefördert wird, bleibt der Umgang mit Bildern entweder vorausgesetzt oder konträr Spezialisten überlassen, z.B. Kunsthistorikern, Wissenschaftlern oder Technischen Redakteuren (vgl. Ballstaedt 1995, S. 63, 1996, S. 192. 200f).

Erst eine semiotische Betrachtung, die darauf hinweist, dass Bilder nicht einfach einen besonderen Typ von visuellen Stimuli darstellen, sondern ikonische Zeichen sind, die ein Sender benutzt, um einem Empfänger etwas zu zeigen, um etwas auszudrücken und etwas zu bewirken, änderte die Perspektive:

Es gibt einige teilweise schon länger existierende theoretische Ansätze und empirische Befunde aus den verschiedensten wissenschaftlichen Disziplinen, welche sukzessive zu einem umfassenden Konzept integriert werden. Die Pädagogik⁵⁰ versteht nicht erst seit Comenius' „Orbis sensualium pictus“ und „Didacta magna“ im 17. Jahrhundert das Bild als Medium sui generis und entwickelt sowohl anwendungsbezogenes Erfahrungswissen sowie verschiedene Entwürfe einer Didaktik der Anschauung. Die Unterrichtswissenschaft⁵¹ ist vor allem daran interessiert, die

⁵⁰ vgl. zusammenfassend Michael (1983).

⁵¹ vgl. zusammenfassend Dallmann und Preibusch (1970).

Effektivität von Bildern in Lernprozessen empirisch zu untersuchen und darüber Merkmale lernförderlicher Bilder zu identifizieren. Auch die Pädagogische Psychologie versucht den Einfluss von Bildern auf Lernerleistungen zu analysieren. Die Wahrnehmungspsychologie betrachtet basale Prozesse des Bilderkennens, u.a. unter entwicklungspsychologischer und kulturvergleichender Perspektive. Die Kunstpsychologie fokussiert Darstellungskonventionen und Erweiterungen des bildlichen Codes sowie das Thema der ästhetischen Effekte von Bildern. Die Semiotik abschließend erforscht Bilder bevorzugt code-analytisch im Vergleich zur Sprache. Grundsätzlich lässt sich festhalten, dass das Verstehen von realen und zeichenhaften Zusammenhängen eine der komplexesten Leistungen des menschlichen Gehirns ist, denn auch wenn das Bildverstehen spontan und automatisch auftritt, ist es ständig mit einer Unzahl von kognitiven Prozessen verbunden (vgl. Ballstaedt 1995, S. 63, 1996, S. 192. 200f; Weidenmann 1988, S. 9f).

Bildspezifika

Eine direkte Gegenüberstellung von deskriptionaler und depiktionaler Vermittlung legt dar, dass zumindest theoretisch jeder Inhalt prinzipiell sprachlich oder bildlich vermittelt werden kann. Dennoch gibt es einige funktionelle Unterschiede zwischen Sprache und Bild in denen sich Bilder als sprachlich unterlegen herausstellen (vgl. Ballstaedt 1996, S. 192; Weidenmann 1988, S. 69):

- Bilder besitzen auf Grund der eingefrorenen Zeit in ihnen keine Analogie zum Verb. Etwaige Äquivalenzen wären, falls überhaupt, das Verbaladjektiv oder Partizip.
- In Bildern gibt es keinen Ausdruck für ein Selbst des Bildproduzenten (in Texten dagegen bspw. in der Form von „Ich“). Bilder sind sogenannte Impersonalien⁵².
- Bilder können für sich allein kein Konzept darstellen. Es kann immer nur das bestimmte Individuum, nie die Gattung repräsentiert werden. Es kann sozusagen nie ein depiktionales Abbild von Kants „Ideen“ geben, es bleibt nur die deskriptionale Repräsentation.
- Bilder können nicht verneinen. Sie stehen quasi im Präsens Indikativ des „so ist es zu diesem Zeitpunkt“.
- Der jeweilige Bildproduzent kann die Verarbeitung seines Produktes nicht explizit kontrollieren – was hingegen einem Textautor möglich ist, indem er den Lesern ausdrückliche Informationen darüber gibt, wie die Informationen zu verarbeiten sind (Informationen z.B. über Aufbau, Zielsetzung, mögliche Fehldeutungen des Textes). Der Textautor kann auf Grund der sequentiellen Struktur von Texten die Verarbeitungsdauer und -abfolge weitgehend steuern. Ein Bild dagegen präsentiert sich simultan und muss den Fokus dem Betrachter überlassen (sofern man von einigen Konventionen – etwa die

⁵² subjektlose Repräsentationen

Bevorzugung von links/rechts sowie oben/unten, also der typisch westlichen Leserichtung – absieht

Die Sprache ist somit eindeutig das Zeichensystem zur Vermittlung abstrakter Begriffe und Zusammenhänge. In ihr gibt es Wörter für Abstrakta, die bildlich nur als Analogien oder Symbole dargestellt werden können. Alsdann lassen sich Aussagen durch Konjunktionen miteinander verknüpfen, so dass in der Sprache begründet und argumentiert werden kann, was bildlich kaum und wenn überhaupt missverständlich möglich ist (vgl. Ballstaedt 1996, S. 194).

Bilder haben jedoch auch gegenüber Texten einige gut erforschte Stärken. Sie können bspw. motivieren und stimulieren, indem sie illustrieren und Anreize zum Hinschauen und zum Weiterbeschäftigen bieten. So wird Lehrmaterial mit Bildern von Lernenden positiver eingeschätzt als reine Texte. Allerdings bedeutet diese Bevorzugung nicht, dass ein Betrachter mit Bildern mehr lernt und behält. Daneben dienen Bilder der Veranschaulichung. Eine Abbildung kann zeigen, wie ein Gegenstand oder ein Detail aussieht. Diese Funktion von Illustrationen als Ersatz für die quasi abwesende Realität ist fundamental, denn sichtbare Merkmale wie Formen, Texturen, Farben sind lediglich stark reduziert sprachlich zu vermitteln. Wenn Lernende Bilder motivierend einschätzen, dann kann diese Einschätzung wahrscheinlich auch durch die Fassbarkeit von Bildern gegenüber den nur wenig unmittelbar die Sinne ansprechenden Texten erklärt werden. Ferner können Bilder die räumliche Orientierung unterstützen. Zwar lassen sich räumliche Verhältnisse und Konfigurationen auch sprachlich vermitteln, gegenüber Bildern erfordert eine Beschreibung jedoch eine größere Verstehensleistung – nämlich einen Moduswechsel – als ein Blick auf eine Abbildung. Desgleichen ist die bildliche Darstellung von Handgriffen, Handbewegungen, Fußstellungen etc. den meisten sprachlichen Beschreibungen überlegen. Abschließend sind Bilder ausnehmend gut geeignet, um Informationen zu verdichten – insbesondere in Form grafischer Visualisierungen. Hierin können Informationen so sehr komprimiert werden, dass sie durch Gruppierungen oder durch ein Bezugssystem holistisch erfassbar sind und damit übergeordnete Strukturen sichtbar werden (vgl. Ballstaedt 1996, S. 192–194; Schnotz/Dutke 2004, S. 176).

Rekapitulation und Präzisierung: Arten von „Bildern“

Der Mensch bedient sich unterschiedlichster Zeichensysteme, um Wissen, Ideen, Konzepte und Vorstellungen zu konservieren und zu kommunizieren. Zeichen sind dabei die sinnlich wahrnehmbaren Stellvertreter für ebendiese Begriffe, Konzepte und Gedanken der Zeichenbenutzer. Die wissenschaftliche Disziplin, die sich mit dem Gebrauch und der Verarbeitung von Zeichen befasst, ist die Psychosemiotik, welche nicht nur die mediale Präsentation von Konzepten via Zeichensystemen untersucht, sondern auch die zu einer mentalen Repräsentation führenden Verarbeitungsprozesse zu analysieren versucht (vgl. Ballstaedt 1996, S. 194).

Es gibt verschiedene Systematisierungen von Zeichensystemen bzw. Repräsentationsformaten – u.a. die Unterscheidung in einfache visuelle Zeichen, Piktogramme, Abbildungen und Visualisierungen. Schnotz (2010) unterscheidet hierbei zusätzlich realistische und logische Bilder (vgl. Schnotz 2010; Niegemann et al. 2004, S. 179).

Zwei sehr unterschiedliche Arten visueller Zeichen sind *ikonische* und *symbolische Zeichen*. Wie bereits angeführt erhalten ikonische Zeichen ihre Bedeutung über konforme perzeptive Merkmale mit dem Signifikat. Dementsprechend werden sie über Entsprechungen in der Wahrnehmung verstanden. Ikonische Einzelzeichen sind jedoch nur schwer definierbar und erhalten ihre Bedeutung erst im konkreten visuellen oder sprachlichen Kontext. Symbolische Zeichen erhalten ihre Bedeutung dagegen ohne konkrete Anbindung an die Realität arbiträr durch Konvention. Sie sind daher nur verständlich, wenn sie speziell gelernt werden. Die Mehrzahl der komplexen visuellen Nachrichten ist semiotisch betrachtet eine Mischform aus verschiedenen Zeichentypen. Dementsprechend stellen die meisten Nachrichten verschiedene Anforderungen an das Verständnis (vgl. Ballstaedt 1996, S. 195f).

Im konkreten Verwendungszusammenhang stehen Zeichen meist nicht allein, sondern zusammengefasst zu Bildern. Dabei kann wie folgt unterschieden werden:

Piktogramme sind einfache, auf einen Blick interpretierbare Bilder, welche einen Begriff aktivieren oder eine Handlung auslösen sollen. Häufig werden in Piktogrammen prägnante ikonische und symbolische Zeichen integriert, deren Verständlichkeit jedoch nicht immer wie gewünscht gegeben ist. Des Öfteren sind Piktogramme zunächst mehrdeutig und werden erst durch den Gebrauchskontext eindeutig (vgl. Ballstaedt 1996, S. 197f).

Abbildungen oder *repräsentionale Bilder* setzen sich aus ikonischen Zeichen zusammen, die einen Realitätsausschnitt in unterschiedlichen Stufen von wahrnehmungsähnlicher Form wiedergeben. Sie werden von Schnotz in die Gruppe der realistischen Bilder gefasst. Vom Produzenten für relevant erachtete Merkmale der optischen Vorlage bleiben in der Abbildung erhalten (Invarianten), gleichzeitig ist diese Abbildung im Vergleich mit der natürlichen Wahrnehmung stets in irgendeiner Weise reduziert: Es können Farben, Oberflächentexturen, Tiefenhinweise etc. fehlen. Eine Abbildung ist somit ein reduzierter Realitätssatz. Dabei können vier kontinuierlich abstrakter werdende Typen unterschieden werden: Fotos, schattierte Zeichnungen, Strichzeichnungen und schematisierte Zeichnungen. Ein *Farb- oder Schwarzweißfoto* ist die realitätsnächste Abbildung. Der Bildausschnitt (von einem bestimmten Standort und damit aus einer Perspektive) legt die Grenzen für die Wahrnehmung fest. Die *schattierte Zeichnung* wird am Zeichentisch oder am Computer erschaffen, um Oberflächeneigenschaften und Beleuchtung zur Steigerung des räumlichen Eindrucks durch Schraffuren und Schattierungen herausheben zu können. Die *Strich-* oder auch *Umrisszeichnung* reduziert die Wahrnehmung auf wenige nüchterne Linien, wobei meist nur Formen und Konturen erhalten bleiben. Sie hat sich in verschiedenen Funktionen als effizient bewährt, da sie auf überflüssige und irritierende Informationen verzichtet. In der *Schemazeichnung* sind charakteristische visuelle Merkmale ggf. bis zur Karikatur hervorgehoben. Ein derartiges Akzentuieren markanter visueller Merkmale scheint

für die Augen auffallend attraktiv zu sein. Empirisch konnte bereits nachgewiesen werden, dass schematisierte Zeichnungen im Vergleich mit anderen Formen der Abbildung rascher erkannt werden. Dies fundiert u.a. Warn- und Sicherheitspiktogramme mit Überzeichnungen bis zur Verzerrung (vgl. Ballstaedt 1996, S. 198f; Niegemann et al. 2004, S. 179f; Schnotz 2010, S. 65f).

Visualisierungen werden hier verstanden als Veranschaulichung nicht wahrnehmbarer Zusammenhänge. Eine Visualisierung transformiert eigentlich nicht-visuelle Relationen in den sichtbaren Bereich. Sie sind damit Schnotz' Gruppe der logischen Bilder zuzuordnen. Visualisierungen werden teilweise als logische oder analytische Bilder bezeichnet, da sie mit dem Sachverhalt, den sie repräsentieren, nur auf struktureller Ebene durch eine Analogierelation verbunden sind. Semiotisch betrachtet können die meisten Visualisierungen als Hybridformen klassifiziert werden, die von verschiedenen Zeichensystemen Gebrauch machen. Notwendig für die Interpretation einer Visualisierung sind die angewandten Darstellungskonventionen, die zumeist gelernt werden müssen und teilweise zu eigenen Fachspezifiken weiter entwickelt wurden. Die wichtigsten Typen von Visualisierungen sind *Karten*, *Schemata*, *Charts* und *Diagramme* (in einigen Theorien auch *Tabellen* – s.u.). *Karten* bilden einen Link von der Abbildung zur Visualisierung, da sie eine Kombination aus Realbild und zusätzlich nicht sichtbaren, konventionalisierten Informationen sind. *Schemata* können als eine auf das Wesentliche reduzierte Zeichnung komplexer Strukturen und Prozesse beschrieben werden. Viele fachspezifische Schemata sind standardisiert hinsichtlich ihrer Darstellungskonventionen. *Charts* bilden qualitative Zusammenhänge zwischen Kategorien oder Aussagen ab. Dabei bilden sie eine Alternative zur Prosa, da ihre Daten wahrscheinlich ohne Defizite, jedoch vernehmlich chaotischer auch in einem Text darstellbar sind. *Diagramme* repräsentieren quantitative Relationen zwischen Variablen (Merkmalen, Eigenschaften, Größen). Hierzu werden nach Schnotz' auch *Tabellen* gezählt. Sie stellen quantitative Daten in einer Matrix aus Zeilen und Spalten dar und können neben exakten Zahlenwerten auch Gruppierungen und Trends abbilden. (vgl. Ballstaedt 1996, S. 199f; Niegemann et al. 2004, S. 180–182; Schnotz 2010, S. 65–67).

Unter Berücksichtigung der vorhergehenden Kapitel ließe sich die Kategorisierung von Lesematerial beispielsweise in einer Übersicht wie der folgenden darstellen, wobei unterschiedliche Farben unterschiedliche Ebenen symbolisieren bzw. gleiche Farben auf gleiche Hierarchielevel verweisen, so dass Teil-Ganzes-Beziehungen erkennbar werden können. Insgesamt muss natürlich eingeschränkt werden, dass es sich dabei um idealisierte Prototypen des jeweiligen Typs handelt und in der konkreten Verwendung in einigen Fällen Abweichungen beobachtet werden können.

Tabelle 3 – Kategorisierung von Lesematerial

symbolische Zeichen (verbal + numerisch)			ikonische Zeichen			
Deskriptionale Repräsentationen		Depiktionale Repräsentationen				
Kontinuierliche Texte	Diskontinuierliche Texte	logische Bilder	analog-repräsentionale Bilder/Abbildungen			
		Visualisierungen ⁵³	schematisierte Zeichnungen (Piktogramme)	Strichzeichnungen	schattierte Zeichnungen	Fotos
		<div>← abstrakt<div></div>konkret →</div>				
Literarische Texte	Listen	Tabellen	Diagramme ⁵⁴			
			Charts			
			Schemata			
			Karten			
Fachtexte/technische Texte (mit jeweils unterschiedlichen Schwerpunkten und Anteilen)						

Zur Erläuterung: der durchgezogene Pfeil bedeutet, dass diese Art der Einteilung zwischen abstrakt und konkret verhältnismäßig eindeutig vorgenommen werden kann und auch die Überschreitung der Grenze ikonisch/symbolisch eindeutig zuordenbar ist. Der gestrichelte Pfeil hingegen gibt lediglich eine häufig zu beobachtende Tendenz an – hierbei überschreitet die Klassifikation „numerisch“ nie die Grenze zwischen symbolisch und ikonisch, da auch numerische Ausdrücke immer Symbole bleiben.

⁵³ Visualisierungen bspw. stellen eine Hybridform zwischen symbolischen und ikonischen Zeichen da (vgl. S. 113), jedoch keine zwischen deskriptionalen Repräsentationen und depiktionalen Repräsentationen, da sie zwar mit konventionalisierten Zeichen arbeiten, jedoch diese nicht primär schriftbasiert sind.

⁵⁴ Diagramme sind eine Unterart von Visualisierungen und schematisierten Zeichnungen und sind im Übergang von logischen Bildern zu analog-repräsentationalen Bildern, die sich im Bereich depiktionaler Repräsentationen befinden (nicht oder kaum schriftbasiert) und können sowohl symbolische Elemente als auch ikonische Elemente enthalten.

Kognitive Verarbeitung und Speicherung von depiktionalen Repräsentationen

Die Beziehung zwischen externer Präsentation und interner Repräsentation von Zeichen ist verhältnismäßig komplex – ein Text kann in mancher Hinsicht in visuelle Vorstellungen umcodiert und Bilder können in anderer Hinsicht verbalisiert werden oder beide verbleiben in ihrem ursprünglichen Format (vgl. Ballstaedt 1996, S. 194).

Bezüglich des Verarbeitungsprozesses (von Bildern wie von Texten) wird in der kognitiven Psychologie meist eine Einteilung in Stufen vorgenommen, selbst wenn eher anzunehmen ist, dass es sich um ein Kontinuum von Prozessen handelt. Das rührt daher, dass eine Betrachtung via abgrenzbare Stufen einfacher vorzunehmen ist auch in dem Bewusstsein einer eingeschränkten Äquivalenz. Es handelt sich hierbei also um eine Modellvorstellung, um sich dem Prozess überhaupt annähern zu können. Zusätzlich laufen die Prozesse auf den Stufen nicht nur seriell ab, sondern simultan und interaktiv verknüpft. Die parallele Verarbeitung der Stufen und innerhalb einer Stufe scheint eine der bemerkenswertesten Leistungen des menschlichen Gehirns zu sein (vgl. Ballstaedt 1995, S. 64, 1996, S. 201; Niegemann et al. 2004, S. 177).

Gemäß dieser Stufenmodellierung beginnt der kognitive Verarbeitungsprozess auf dem Level der präattentive Verarbeitung mit der perzeptiven Encodierung. Zunächst ist festzuhalten, dass ein kurzer (100 msec) Blick auf ein Bild ausreicht, um ein Objekt, eine Szenerie oder eine Handlung wahrzunehmen, zu erkennen und zu kategorisieren. Ermöglicht wird das mit Hilfe angeborener Verarbeitungsoperationen, welche automatisiert und simultan eine perzeptuelle Disposition formen. Anteil an diesem Prozess haben entscheidend die sogenannten Gestaltgesetze und die Mustererkennung visueller Schemata. Diese Decodierprozesse verlaufen parallel und beinhalten automatisierte visuelle Routinen. Dazu sind sie primär datengeleitet und folglich relativ unabhängig vom Vorwissen sowie den Intentionen des Individuums (vgl. Ballstaedt 1995, S. 64; Schnotz 2010, S. 70f)

Biologisch betrachtet scheinen in den Neuronennetzen des menschlichen Gehirns Organisationsregeln – die oben bereits genannten Gestaltgesetze – einprogrammiert zu sein, welche singuläre Punkte, Striche und Flächen zu Gestalten oder Figuren vor einem Hinter- oder Untergrund konstruieren und/oder ordnen (Figur-Grund-Einteilung) und anschließend entsprechend der Gestaltfaktoren gruppieren. Das Zusammenwirken verschiedener Subprozesse führt schließlich zu einer kohärenten und gegliederten Wahrnehmung. Jedoch ist die Analyse und Forschung bezüglich des visuellen Systems noch nicht in der Lage, den Ablauf aller Prozesse vollständig zu erklären (vgl. Ballstaedt 1996, S. 201; Niegemann et al. 2004, S. 177f).

Auch wenn die intuitive und unwillkürliche Organisation des Sehfeldes simpel erscheint, ist ihr Zustandekommen bisher für die Neuropsychologen auch heute noch rätselhaft. Es ist zwar bekannt, dass das Gehirn visuelle Sinnesdaten in diversen separaten Arealen synchron auswertet: ein Areal analysiert die Formen, eines die Farben, ein weiteres die räumlichen Beziehungen, ein anders die Bewegungen und eines speziell menschliche Gesichter. Problematisch, weil unaufgeklärt, bleibt allerdings, wie diese getrennten Analysen zu einer kohärenten Interpretation

synthetisiert werden (vgl. Ballstaedt 1996, S. 203). Bekannt ist lediglich, dass die unterschiedlichen Gestaftgesetze nach einem Prägnanzprinzip zusammen wirken. Es sagt aus, dass das menschliche Gehirn aus der Menge der Sinnesdaten kontinuierlich die einfachste mögliche Anordnung herausrechnet. Da die Funktion dieser Stufe eine rasche räumlichen Orientierung ist, deren biologischer Nutzen offenkundig ist und sich die Entwicklung des Gehirns in Interaktion mit der Umwelt vollzieht, kann im Allgemeinen davon ausgegangen werden, dass die Wahrnehmung natürlicher Umgebungen verlässlich und nur selten irreführend ist – so kommen optische Täuschungen in der Natur für gewöhnlich nicht vor (vgl. Ballstaedt 1995, S. 64, 1996, S. 201).

Die Verarbeitung auf dieser Stufe wird weitestgehend als unbewusst und willentlich nicht beeinflussbar eingestuft. Dass diese Stufe des Verarbeitungsprozesses überhaupt existiert, wird nur bemerkt, sofern keine prägnante und eindeutige perzeptuelle Organisation unmittelbar zustande kommt und durch bewusste Strategien ergänzt werden muss. Ein Bildtyp, dessen Verstehen im Optimalfall auf dieser Stufe der Verarbeitung vollständig vollzogen werden kann, sind Piktogramme (vgl. Ballstaedt 1995, S. 64)

Auf die präattentiv konstruierte perzeptuelle Repräsentation folgen attentive Ablese- bzw. Inspektionsprozesse zur semantischen Encodierung, um die Repräsentation zu interpretieren bzw. um ihr bestimmte Informationen zu entnehmen. Diese Prozesse laufen mit hoher Wahrscheinlichkeit eher seriell ab und sind sowohl daten- als auch konzeptgeleitet. Sie beinhalten also ein Wechselspiel von aufsteigender und absteigender Aktivierung kognitiver Schemata. Sie werden dementsprechend im Gegensatz zu den präattentiven Prozessen sowohl vom Vorwissen als auch von den Zielsetzungen und Intentionen des Individuums beeinflusst (vgl. Schnotz 2010, S. 71).

Auf dieser Stufe beginnen Detailauswertungen, deren beobachtbarer und messbarer Indikator die sakkadischen Augenbewegungen sind. Diese sind im Gegensatz zu den andern Prozessen eindeutig seriell: Ein Bildausschnitt nach dem anderen wird fixiert und damit in die Stelle des schärfsten Sehens auf der Retina gebracht. Die Sakkaden und Fixationen lassen sich messen und das entstandene Muster gibt Hinweise auf die Aufteilung von Aufmerksamkeit. In einer Sekunde sind durchschnittlich drei Fixationen möglich. Die Dauer und Häufigkeit der Zuwendung der Augen können so als ein Indikator für visuelles Interesse und Konzentration dienen. Die Augenbewegungen sind nur zu einem kleinen Teil reflektorisch, der größere Anteil ist bewusst kontrolliert und durch Interessen, Vorwissen, Intentionen oder Aufgaben gesteuert, d.h. der Betrachter entscheidet bewusst, was er genauer, d.h. öfter und länger anschaut. Die Funktion dieser Verarbeitungsstufe ist die aktive Exploration der visuellen Umgebung. Dazu hat der Mensch scheinbar im Laufe der Evolution und Kultur sogar domainspezifische visuelle Routinen entwickelt (vgl. Ballstaedt 1995, S. 64f, 1996, S. 203; Niegemann et al. 2004, S. 177).

Es wird angenommen, dass der Betrachter bei einer Fixation die visuellen Informationen zunächst kategorisiert und teilweise auch in eine verbalisierte Form bringt. So kann eine (stück-

weise) sprachliche Kopie oder eine mehr oder weniger ausführliche Beschreibung der depiktionalen Repräsentation entstehen. Während einer Fixation entscheidet der Betrachter mit Hilfe von Vorwissen und beeinflusst durch die gebotenen Reize, welches Bildareal als nächstes fokussiert wird (vgl. Ballstaedt 1996, S. 203).

Dieser Prozess wird durch zwei stark kooperierende Kontrollmechanismen geleitet: Erstens der reflektorischen Steuerung – diese wird durch aufmerksamkeitsregende Merkmale oder Anreize der visuellen Vorlage ausgelöst. Es konnte beobachtet werden, dass ungewöhnlichen und überraschenden Informationen (z.B. grelle Farben, seltene Schriften, unerwartete Objekte) zuerst betrachtet werden und Areale mit großer Informationsdichte länger und intensiver studiert werden. Der zweite Kontrollmechanismus ist die willentliche Steuerung. Die Detailauswertung einer depiktionalen Repräsentation ist eine aktive Tätigkeit des Betrachters, die je nach Interesse, Vorwissen und Intention anders ausfallen kann. Die freie Bildbetrachtung wird gesteuert durch Schemata, die die Selektion von Zielarealen für die Blicksprünge beeinflussen. Unter einem Schema wird in diesem Kontext begriffliches oder visuelles Wissen über typische Zusammenhänge eines Realitätsbereichs verstanden (vgl. Kapitel 3). Ein Schema wird an der Erwartung erkennbar, an einer speziellen Stelle in einem Bild etwas Bestimmtes zu sehen (vgl. Ballstaedt 1996, S. 203f).

Im Verlauf der aufmerksamen Verarbeitung werden parallel sprachliche Prozesse initiiert, die ein Objekt oder eine Person benennen und die Beziehungen zwischen ihnen beschreiben. Durch die Durchmusterung der Vorlage und die damit verbundene Aktivierung und Verbalisierung von Konzepten konstruiert der Betrachter sein semantisches Verständnis der Abbildung (vgl. Ballstaedt 1996, S. 205; Niegemann et al. 2004, S. 178).

Die folgende Ebene wird als elaborative Verarbeitung bezeichnet. Diese dritte Ebene ist indifferent zu bestimmen, möglicherweise kann diese ebenfalls als Vertiefung der zweiten Ebene angesehen werden. Auf dieser Ebene werden über die Abbildung hinausgehende und von der Abbildung ausgelöste Schlussfolgerungen, Assoziationen und Vorstellungen, welche durch vorhandenes Wissen (Schemata) freigesetzt werden, verortet. Auf dieser Stufe wird vom Betrachter typischerweise versucht zu eruieren, warum der Produzent dieses Bild in dieser Form (und ggf. in diesem Kontext) präsentiert. Teilweise werden diese Prozesse als indikatorisches Verstehen bezeichnet. Als Funktion dieser Verarbeitungsstufe wird die multiple Integration in vorhandene Wissensbestände angegeben (vgl. Ballstaedt 1995, S. 65; Niegemann et al. 2004, S. 178).

Weidenmann (1988) unterscheidet von der elaborative Ebene noch ein ökologisches Bildverstehen, bei dem der Betrachter lediglich erkennt, was auf dem Bild dargestellt ist und in ein davon zu unterscheidendes indikatorisches Bildverstehen. Bei diesem geht er davon aus, dass beim Betrachter auch Überlegungen bezüglich der Absichten des Bildproduzenten über die Darstellung des Sachverhalts auf genau diese Weise bzw. bezüglich der Entscheidung gerade dieses Bild zu präsentieren stattfinden. Das Bild wird an dieser Stelle gewissermaßen als Indikator für eine bestimmte Mitteilungsabsicht angesehen und entsprechend analysiert (vgl. Weidenmann

1988, S. 168). Damit repräsentieren ökologisches und indikatorisches Bildverstehen jedoch derart unterschiedliche Verstehentiefen, dass eigentlich nicht mehr von derselben Stufe gesprochen werden kann (vgl. Schnotz 2010, S. 71).

Mit der Anfertigung und/oder Verwendung eines Bildes vollzieht der Produzent bzw. Redakteur auch eine Handlung, einen Zeigeakt. Im Gegensatz zur Sprache, in der Kundgabe und Appell direkt ausgedrückt werden können, lassen sich deren Äquivalente eines Bildes nicht direkt darstellen – es kann nicht unmittelbar zeigen, wozu es dienen soll oder anders formuliert: Bilder können sich nicht selbst kommentieren (vgl. Ballstaedt 1996, S. 205).

Der Betrachter muss diese kommunikative Funktion erst aus Gestaltungsmerkmalen erschließen. Daher leuchtet ein, dass das intentionale Verstehen somit verhältnismäßig störanfällig sein kann. Es setzt ein indikatorisches Sehen voraus, das die Art und Weise der Gestaltung als Hinweis auf die Absichten des Bildproduzenten auswertet (vgl. Ballstaedt 1996, S. 205; Niegemann et al. 2004, S. 178).

Theoretisch ist denkbar, dass die elaborative Stufe unter Umständen eine jahrelange Auseinandersetzung mit einem Bild umfassen kann. Es kann kein definitorisches Ende des Prozesses gegeben werden – analog zum hermeneutischen Zirkel der (Inter-)Textkommunikation, der auch potentiell unendlich verlaufen kann. Weiterhin belegen empirische Befunde, dass diese Stufe der Bildverarbeitung und des -verstehens nicht selbstverständlich ist, sondern gelernt und trainiert werden muss (Schmidt et al. 1989 zitiert nach Ballstaedt 1995, S. 65).

Einige Forscher schreiben dem Bildverstehen noch eine weitere Ebene der rekonstruktiven Verarbeitung zu. Hier überschneiden sich Verarbeiten, Speichern und Erinnern. Wird ein Bild nicht als komplexes Ganzes im Gedächtnis abgespeichert und später wieder abgerufen, können wesentliche Objekte des Bildes wie die strukturelle Anordnung, die farbliche Gestaltung oder die Aussage des Bildes fehlen. Nur was zuvor eingeprägt wurde, kann später überhaupt rekonstruiert werden. Da es schwierig bis unmöglich für gewöhnliche Menschen ist, sich alle Details eines Bildes einzuprägen, wirken die erstellten Rekonstruktionen häufig merkmalsärmer als ihre Originale. Interessant ist aber, dass trotz dieser „Fehlerhaftigkeit“ selbst nach Tagen ein ehemals präsentiertes Bild noch richtig wiedererkannt werden kann (vgl. Niegemann et al. 2004, S. 179).

Die beschriebenen Stufen finden sich bisher in allen theoretischen Ansätzen zur Bildverarbeitung. Auch wenn es sich um ein Kontinuum an Verarbeitungsprozessen statt um Stufen handelt, kann doch festgestellt werden, dass sie durch einen zunehmenden Verarbeitungsaufwand und zunehmende Verarbeitungstiefe gekennzeichnet sind und nicht immer alle erfolgen müssen – es kann ggf. nur ein Blick auf ein Bild geworfen werden (präattentive Verarbeitung), ein Bild kann ausführlich exploriert werden (attentive Verarbeitung) und schließlich können weiterführende Gedanken investiert werden (elaborative Verarbeitung). Von Stufe zu Stufe nimmt dabei das Kontingent der top-down-Verarbeitung zu (vgl. Ballstaedt 1995, S. 65). Auch international wurde bisher (Stand: 2015) nichts Gegenteiliges veröffentlicht.

Auch wenn logische Bilder bzw. Diagramme einen Sachverhalt nicht auf Grund von Ähnlichkeit, sondern auf Grund von abstrakteren strukturellen Gemeinsamkeiten repräsentieren muss beim Verstehen eines Diagramms nicht von grundlegend anderen Prozessen ausgegangen werden. Es ist somit weithin akzeptiert, dass der Betrachter ein mentales Modell des dargestellten Sachverhalts konstruiert. Die dabei stattfindenden subsemantischen, präattentiven Prozesse setzen sich sowohl aus der Diskrimination und Identifikation sowie andererseits auch aus der Gruppierung grafischer Komponenten – also von Punkten, Linien und Flächen – entsprechend den sog. Gestaltgesetzen (s.o.) zusammen und führen so zur Wahrnehmung einer entsprechenden grafischen Konfiguration (vgl. Schnotz 2010, S. 72).

Nach Schnotz (2010) besteht die semantische, attentive Verarbeitung beim Verstehen von Diagrammen aus einer konzeptgeleiteten Analyse der wahrgenommenen grafischen Konfiguration. Hierbei werden einige visuell-räumliche Relationen und Attribute abgelesen, sodann semantisch interpretiert und vermutlich in Form von Propositionen fixiert. Diese werden abschließend der bereits vorhandenen propositionalen mentalen Repräsentation hinzugefügt. Da Diagramme im Gegensatz zu realistischen Bildern keine perzeptuelle Ähnlichkeit mit dem repräsentierten Gegenstand besitzen, kann der Betrachter bei der Interpretation logischer Bilder nicht auf kognitive Schemata der alltäglichen Wahrnehmung zurückgreifen. Somit ist die Fähigkeit zum Verstehen von Diagrammen eine spezifische Kulturtechnik, die speziell erlernt werden muss. Dabei werden spezielle Grafikschemata konstruiert, mit deren Hilfe an den grafischen Konfigurationen eines Diagramms bestimmte Informationen abgelesen werden können. Verfügt ein Betrachter nicht über entsprechende Schemata, so kann er wahrscheinlich einem Diagramm die jeweiligen Informationen nicht entnehmen und es kaum korrekt interpretieren bzw. stößt in diesem Prozess auf erhebliche Schwierigkeiten. Dies wird vor allem daran deutlich, wenn er nicht in der Lage ist einer solchen grafischen Darstellung relevante Informationen über komplexe Zusammenhänge zu entnehmen (vgl. Schnotz 2010, S. 72).

Empirische Untersuchungen zeigen zudem, dass Experten bzw. Betrachter mit höherem Vorwissen besser in der Lage sind, übergreifende visuelle Muster zu suchen und zu erkennen, wobei Betrachter mit geringerem Vorwissen vornehmlich eine lokal begrenzte Suche nach Einzelinformationen vornehmen. Somit kann vermutet werden, dass diese Experten oder Individuen mit höheren Verständnisvoraussetzungen über elaboriertere und stärker hierarchisch strukturierte Grafikschemata verfügen (vgl. Schnotz 2010, S. 72).

Summarisch komprimiert heißt das: für das Verstehen depiktionaler Strukturen muss der Betrachter nachfolgende mentale Repräsentationen generieren (vgl. Schnotz/Dutke 2004, S. 73):

1. eine perzeptuelle Repräsentation des Bildes bzw. Diagramms,
2. ein mentales Modell der dargestellten Sachlage,
3. eine propositionale mentale Repräsentation der dargestellten Sachlage,
4. eine mentale Repräsentation der mit der depiktionalen Repräsentation verfolgten Kommunikationsabsicht,

5. eine mentale Repräsentation des Bild- bzw. Diagrammgenres.

Bei einem diesbezüglichen Umgang mit Depiktionen stellen sich sodann mannigfache Anforderungen auf drei heteromorphen Ebenen (vgl. Lachmeyer 2008, S. 31f): Die Ebene der Informationsentnahme, hierbei geht es um Prozesse der Informationsverarbeitung – d.h. inhaltliche Interpretation von Diagramminformationen. Neben der schlichten Identifikation von Daten kann die Registrierung als gesonderter Bereich der Ebene der Informationsentnahme gesehen werden (vgl. Bertin 1974). Des Weiteren die Konstruktionsebene, welche es erfordert aus einer gegebenen Anzahl an Daten oder Informationen, eine Diagrammaussage komplett eigenständig zu konstruieren. Und drittens die Integrationsebene. Auf dieser Ebene wird vom Leser des Diagramms verlangt, die aus dem Diagramm extrahierten Hinweise mit anderen Daten wie bspw. dem jeweiligen Vorwissen oder einem zugehörigen Fließtext in Bezug zu setzen (vgl. Lachmeyer 2008, S. 43).

Eine unter Psychologen noch umstrittene Frage ist, wie genau das menschliche Gehirn visuelle Informationen aufbewahrt. Seit längerem gibt es die Erkenntnis, dass visuelle Präsentationen bemerkenswert gut memoriert werden. In Experimenten bekamen Versuchspersonen bis zu 10 000 einfache Bilder kurz präsentiert. Noch nach Tagen erkannten sie über 90% unter anderen, nicht gebotenen Bildern wieder, und selbst nach einem Jahr haften Bilder noch gut im Gedächtnis (vgl. Standing 1973; Standing et al. 1970; Hirst 1990; zitiert nach Ballstaedt 1996). Dabei ist das menschliche Gedächtnis kein einheitliches Speichersystem, sondern segmentiert in funktionale Subsysteme (vgl. Engelkamp 1990). Es wird vermutet, dass komplexe Informationen multipel, d. h. in unterschiedlichen Repräsentationen und in verschiedenen Subsystemen im Gehirn gespeichert werden. Was genau aufbewahrt wird, scheint eine Art Verarbeitungsprotokoll, d.h. eine konzeptuelle Chronik von Auswertungsprozessen, die während der Bildverarbeitung abgelaufen sind, zu sein. Erinnern ist deshalb wie oben bereits angeführt nicht die simple Entnahme eines fertigen Erzeugnisses aus einem Speicher, sondern eine erneute Konstruktion (→ Rekonstruktion) auf Grund vorhandener Prozessdaten. Es können wahrscheinlich folgende Gedächtnisrepräsentationen unterschieden werden:

Visuelle Repräsentation: Durch neuropsychologische Erkenntnisse scheint gesichert zu sein, dass es ein eigenes visuelles Langzeitgedächtnis gibt. Es bringt statische wie bewegte visuelle Erinnerungen und Vorstellungen hervor, die zumeist ausdrucksleerer und schlichter als die Originalwahrnehmung sind (vgl. Ballstaedt 1996, S. 207).

Konzeptuelle Repräsentation: Die bei der Detailauswertung produzierte Teilversprachlichung eines Bildes ist als begrifflich-sprachliche Charakteristik gespeichert. Explizit verbalisierte Inhalte werden länger behalten als nicht verbalisierte. Dies lässt sich vermutlich auf die sodann vorhandene doppelte Repräsentation als visuelle und sprachliche Information zurückführen (vgl. Ballstaedt 1996, S. 207).

Motorische Repräsentation: Abfolgen von Handgriffen und Handlungen (vgl. prozedurales Wissen) werden vermutlich als motorische Programme repräsentiert, die sich teilweise dem Bewusstsein und der Verbalisierung entziehen. Die Aktivierung eines motorischen Programms beinhaltet zumeist also auch die Aktivierung der dazugehörigen dynamischen visuellen Repräsentation. Es ist somit also akkurater eine sensumotorische Repräsentation (vgl. Engelkamp 1990; Ballstaedt 1996, S. 208).

Mentale Modelle (vgl. 0) sollen die bisher aufgeführten Repräsentationstypen integrieren. In der Forschung nicht unumstritten ist, ob diese im Eigentlichen überhaupt eine autarke Repräsentation darstellen. Besonders im technischen Bereich wird eine Repräsentation von komplexen Zusammenhängen als mentalen Modelle, d.h. als eine analoge und dynamische mentale Repräsentation von Konstellationen, Gegebenheiten und Ereignissen, jedoch stark postuliert. Mit einem mentalen Modell können Abläufe sodann in der Vorstellung simuliert werden. Ballstaedt führt dazu plausibel aus: „Da ein mentales Modell somit Anschaulichkeit und räumliche Orientierung voraussetzt, sind Bilder besonders zum Aufbau derartig anspruchsvoller Repräsentationen geeignet.“ (Ballstaedt 1996, S. 208)

Für die Annahme eines mentalen Modells auch beim Bildverstehen spricht, dass eine derartige integrierte und trotzdem multiple Aufbewahrung von Informationen in ungleichartigen Repräsentationsformaten den Vorteil hat, dass jeweils auf das Format zurückgegriffen werden kann, welches für die konkrete Problemlösung optimal scheint (vgl. Ballstaedt 1996, S. 208).

Modelle der Integration von Text- und Bildverarbeitung

Bereits in der Theorie mentaler Modelle ist der Grundstein gelegt für eine gemeinsame Betrachtung von Text und Bild. Die wesentliche Annahme der Theorie ist, wie bereits ausgeführt, dass bei Textverstehensprozessen nichtsprachliche Repräsentationen konstruiert werden, welche im Prinzip von derselben Art sind wie die Repräsentationen, die Personen bei nichtsprachlichen kognitiven Tätigkeiten konstruieren. Ausschlaggebend ist also die Relation zwischen den beim Textverstehen gebildeten Repräsentationen und den Repräsentationen, die bei der Wahrnehmung, der Vorstellung, der Erinnerung, der Handlungsplanung oder der Lösung nicht-sprachlicher Probleme gebildet werden (vgl. Kelter 1995, S. 73).

Glenberg/Langston (1992) legen eine erste Konzeption einer Mentale-Modell-Theorie des „Text“verstehens vor, welche sowohl Theorien als auch empirische Befunde der Forschung zur nicht-sprachlichen Kognition integriert. Dabei wird die bereits etablierte Mentale-Modell-Theorie des Textverstehens mit einer Arbeitsgedächtnis-Theorie von Baddeley (1986) bzw. Baddeley/Logie (1999) verbunden. In dieser integrativen Theorie wird besonders das sogenannten Visuo-Spatial-Sketchpad (VSSP) berücksichtigt und als bedeutend eingestuft. Es ist gemäß der Theorie von Baddeley ein System für die kurzzeitige Repräsentation perzeptuell oder anders vermittelter räumlicher Informationen. Sie sind der Ansicht, dass bei der Textrezeption jede im Text erwähnte Entität durch ein repräsentationales Element im VSSP abgebildet wird. Solche Elemente im VSSP beinhalten selbst keine Information über die Eigenschaften der entsprechenden Entität, sondern rekurren nur auf die entsprechenden Informationen, welche in anderen

unterschiedlichen mentalen Systemen deponiert sind – zum Beispiel auf die Informationen über Form und Gestalt der Entität. Bedeutsam ist an dieser Stelle vor allem die Organisation der repräsentationalen Elemente im VSSP. Diese sind im VSSP für gewöhnlich so angeordnet, dass sie die räumlichen Relationen zwischen den entsprechenden Elementen in dem realen (text- oder bildvermittelten) Sachverhalt analog repräsentieren. Wenn also in dem geschilderten Sachverhalt Element A näher bei Element B als bei Element C liegt, werden die entsprechenden repräsentationalen Entitäten im VSSP ebenso organisiert, d.h. die Distanz A-B ist geringer als die Distanz A-C. Unter bestimmten Umständen kann das VSSP auch eingesetzt werden, um andere Relationen wie beispielsweise chronologische/temporale Relationen zwischen Ereignissen analog darzustellen. Dieses ist der Fall, wenn die Relationen dem Rezipienten bei dem präsentierten Text besonders relevant anmuten (vgl. Glenberg et al. 1994). Bedeutend an dem Theorieansatz sind vor allem zwei Annahmen. Zum einen hat die Hypothese, dass verschiedene Arten von Informationen in verschiedenen kognitiven Systemen repräsentiert werden, entscheidende Implikationen für die Forschung. Gemäß dieser Annahme ist sodann eine Sachverhaltsrepräsentation nicht einfach eine einzelne Repräsentation, sondern ein Komplex von Repräsentationen, der separat für sich untersucht werden müsste. Zum anderen ist aufschlussreich, dass die zentrale Rolle für das Textverstehen einem kognitiven System zugeschrieben wird, welches ansonsten als System für die Repräsentation lokaler Information bei der Perzeption und Vorstellung eingeordnet wird (vgl. Kelter 1995, S. 74).

Daraus ergeben sich folgende Konsequenzen: Normalerweise gelten jene Repräsentationen des VSSP als mentale Vorstellungen, welche nicht unmittelbar zuvor aufgenommene Stimuli repräsentieren, sondern durch den Input andersartiger bspw. sprachlicher oder nichtsprachlicher Informationen konstruiert wurden (vgl. Baddeley 1986, S. 109f).

In dem Theorieansatz von Glenberg/Langston (1992) ist die Vermutung enthalten, dass ein wesentlicher Teil des Textverstehensprozesses zum Teil aus der Konstruktion räumlicher Vorstellungen besteht. Damit könnte angenommen werden, dass die beim Textverstehen konstruierten mentalen Modelle theoretisch dieselben Eigenschaften aufweisen wie räumliche Vorstellungen und für sie auch dieselben Axiome der Verarbeitung gelten könnten. Sodann können die Ergebnisse der Analysen zu lokalen Dispositionen genutzt werden, um vorliegende Untersuchungen zu Textverstehensprozessen in einem größeren theoretischen Zusammenhang zu interpretieren und um fundierte Thesen über die bei der Textverarbeitung ablaufenden Prozesse zu formulieren (vgl. Kelter 1995, S. 75). Eine Betrachtung der Beziehungen zwischen Textverstehensprozessen und räumlichen Vorstellungen ist auch auf Grund der Tatsache besonders interessant, dass diese Vorstellungen eine enge Interaktion mit der Wahrnehmung aufweisen können und sich somit ggf. die Möglichkeit eröffnen kann, auch die Beziehungen zwischen sprachlich vermittelten und perzeptuell vermittelten Informationen genauer zu untersuchen (vgl. Kelter 1995, S. 80).

Eine gemeinsame Betrachtung und Parallelisierung der Prozesse des Bild- und Textverstehens ist der erste Schritt zu einer echten interaktionistischen Theorie bei der die gemeinsame, reziproke Präsentation und Verarbeitung von Bild und Text im Fokus steht.

Es kann konstatiert werden, wie die Ergebnisse aus den Kapiteln 0 und 0 ebenfalls darlegen, dass Bilder oder Texte selten gänzlich isoliert betrachtet werden. Stattdessen sind Bilder meist mit deskriptionalen Strukturen kombiniert und Texte werden durch depiktionale Repräsentationen ergänzt.⁵⁵ Ferner bewegen sich logische Bilder ohnehin in diesem Spannungsfeld. Bei einer Verarbeitung der gebotenen Informationen von Text und Bild müssen diese sprachlichen und bildlichen Mitteilungen wechselseitig analysiert und integriert werden. Eine Zusammenführung kann dann als gelungen bezeichnet werden, wenn sich die Bildanalyse und das Textverstehen beiderseits begünstigen (vgl. Ballstaedt 1996, S. 221).

Auch wenn nicht ausgeschlossen wird, dass anschauliches Denken zentral an basalen Verarbeitungsprozessen beteiligt ist, wird doch davon ausgegangen, dass die höheren Verarbeitungsprozesse weitgehend an Sprache gebunden sind (vgl. Ballstaedt 1995, S. 67). Untersuchungen bezüglich der Frage, ob bei der Betrachtung von Bildern sprachliche Prozesse tatsächlich immer beteiligt sind, sind durch Introspektion schwierig zu erforschen und zu beantworten. Ballstaedt (1995) führt dazu lapidar aus: „Manchmal kommen einem Worte oder Sätze in den Kopf, manchmal offensichtlich nicht“ (Ballstaedt 1995, S. 66). Was Untersuchungen (Nelson et al. 1976; Nelson/Brooks 1973) zeigen konnten, ist dass es einen deutlichen Effekt bestimmter Einflussfaktoren der jeweils unabhängigen Variablen gibt. Die Forscher untersuchen in einem sogenannten Paarlernen-Experiment zwei unabhängigen Variablen in ihrer Wirkung auf die Lernleistung (als abhängige Variable). Die Homonymie⁵⁶ von Wörtern, dazu benutzen sie Wörter einer Liste, welche zwar semantisch unverbunden waren, aber homonym waren. Die Wörter der anderen Liste waren lautlich völlig verschieden. Zweitens untersuchten sie die Darbietungsform. Es wurden Wort-Wort-Paare und Bild-Wort-Paare benutzt. Die Ergebnisse gestalten sich wie bereits angedeutet folgendermaßen: Eine große Homonymie führt bei der Wort-Wort-Bedingung zu einer signifikant schlechteren Lernleistung gegenüber den Wörtern mit geringer Lautähnlichkeit. Dieser Effekt kann durch Interferenzen zwischen den Reizwörtern erklärt werden. Visuelle Informationen konstituieren Konzeptionen, im Gegensatz dazu scheinen Wörter jedoch lediglich Platzhalter oder Stellvertreter für Konzepte zu sein. Es gibt weitere empirische Untersuchungen mit anderen Methoden, die bezüglich dieses Befunds als weitere Bestätigung dienen können. So konnte in psycholinguistischen Untersuchungen festgestellt werden, dass sich auch die Rekognitionsleistung signifikant steigert, wenn eine Aufgabe explizit die Verbalisierung komplexer Bilder verlangt (Bartlett et al. 1980). Im umgekehrten Fall, dass die Unterdrückung eines Verbalisierungsprozesses durch ablenkende Aufgaben – bspw. während der Bildbetrachtung in Dreierschritten rückwärts zu zählen – die Rekognitionsleistung beeinträchtigt (Loftus 1972; Loftus/Kallman 1979) konnte ebenfalls nachgewiesen werden (vgl. Ballstaedt 1995, S. 67).

Eine Reihe von Untersuchungen hat weiterhin auch aufgezeigt, dass ein sprachlicher Kontext die Verarbeitung und Auslegung eines Bildes beeinflussen kann. Das Bildverstehen kann sprachlich bspw. dadurch angeleitet werden, dass vorher eine Wahrnehmungseinstellung erzeugt

⁵⁵ Wenngleich insbesondere der geisteswissenschaftliche Gebrauch ein anderer ist.

⁵⁶ Laut-/Klangähnlichkeit

wird, dadurch, dass während der Betrachtung die visuelle Aufmerksamkeit gelenkt wird oder dadurch, dass nachher die elaborative Verarbeitung beeinflusst wird. Es können drei Möglichkeiten der Text-Bild-Interaktion unterschieden werden (vgl. Ballstaedt 1995, S. 68):

1. eine zusätzliche Erklärung/Anmerkung kann die Einstellung des Betrachters zur Abbildung verändern (z.B. der Hinweis, ob ein Bild im Urlaubs- oder im Werbekontext entstanden ist; ob ein Bild eine Vase oder zwei Gesichter darstellt...),
2. eine sprachliche Anleitung kann die Augen auf bestimmte Bilddetails (Akzentuierung) lenken oder bestimmte Blickpfade nahelegen,
3. durch sprachliche Bezeichnungen kann eine Kategorisierung der Perzepte vorgegeben werden.

Entscheidend für die Behaltensleistung scheint nach Loftus (1972) die Anzahl der Fixationen zu sein. Eine Erklärung dafür könnte sein, dass während einer Fixation nicht nur die visuellen Informationen wahrgenommen und analysiert werden, sondern sich auch Prozesse der Kategorisierung und der Verbalisierung vollziehen. Aus dem entstehenden Fixationsmuster lassen sich zumeist die Inhalte einer freien Reproduktion vorhersagen (vgl. Ballstaedt 1995, S. 67f).

Wie in Kapitel 0 ausgeführt, ist es wahrscheinlich, dass Bilder multimodal mental repräsentiert werden: Einerseits als depiktionale Entitäten in einem visuellen Langzeitgedächtnis, andererseits als teilweise sprachliche (propositionale) Kopie und weiterhin als rein konzeptuelle Repräsentation, die keine analoge oder propositionale Entsprechung mehr hat (Gefühle, Geruch). Welches Repräsentationsformat genutzt wird, hängt von den Ansprüchen der Aufgabe ab. So dann sind Spuren von Verarbeitungsprozessen auf verschiedenen Stufen repräsentiert, respektive dem, wie und in welchem Kontext sich der Betrachter mit dem Bild beschäftigt hat. Die verschiedenen vorhandenen Repräsentationen können in Abhängigkeit von der Aufgabenstellung weiter verarbeitet oder eingesetzt werden (vgl. Ballstaedt 1995, S. 68).

Auch wenn die meisten Kognitionspsychologen heute davon ausgehen, dass sowohl beim Verstehen von Texten als auch beim Verstehen von Bildern und Diagrammen multiple mentale Repräsentationen konstruiert werden (Schnitz 2010, S. 67), ist noch umstritten, wie diese Konstruktion erfolgt, und welche „Kanäle“ und neuronale Verarbeitungsmodule genau genutzt werden.

Paivio (1969; 1971; 1990), Clark/Paivio (1991) und Sadoski/Paivio (2001; 2004) nehmen in ihrer Dual-Coding-Theory (duale Codierungstheorie, im Folgenden DCT) zwei getrennte, aber interagierende Systeme zur Verarbeitung von sprachlichen und bildlichen Informationen an – einerseits ein verbales System zur Verarbeitung sprachlicher Daten und ein imaginales System zur Verarbeitung bildhafter Daten. Sätze oder Texte werden demnach meist nur im verbalen System encodiert, während Bilder grundsätzlich imaginal und verbal encodiert werden. Dieses Konzept, nach dem unterschiedliche Informationsarten durch verschiedene Kanäle rezipiert

werden, hat eine lange Tradition in der Kognitionspsychologie (vgl. Niedermann et al. 2004, S. 188).

Dass sich die menschliche Informationsverarbeitung danach unterscheidet, ob auditive oder visuelle Informationen erfasst werden sollen ist eine zentrale These der Theorie. Um diese präsentierten Informationen zu verarbeiten, werden jeweils unterschiedliche Kanäle angesprochen: für verbale Daten der auditiv/verbale Kanal und für visuelle Daten der visuell/bildhaft/nonverbale Kanal (vgl. Sadoski/Paivio 2004, S. 3).

Letzterer wird häufig als „imagery system/code“ bezeichnet, da seine Funktionen die Erzeugung, Analyse und Transformation von mentalen Bildern umfasst. Jeder „Code“ hat seine eigenen charakteristischen Einheiten und hierarchische Organisation. Zusammen bilden die beiden Systeme die Kenntnisse der Sprache und das Wissen der Welt ab. Diese beiden psychischen Codes und die fünf Sinne sind in der DCT orthogonal angeordnet. Dies bedeutet, dass die beiden Codes jeweils Teilmengen mentaler Repräsentationen besitzen, welche qualitativ verschieden sind auf Grund der unterschiedlichen Sinnesempfindungen, aus denen sie gewonnen werden. Da Sinnessysteme an motorische Feedbacksysteme in der Wahrnehmung – bspw. Augenbewegungen, aktive Berührung – gekoppelt sind, haben auch diese Teilmengen sensomotorische Qualitäten. Menschen entwickeln visuelle Darstellungen im Sprachcode für verbalen Einheiten, die, wie Buchstaben, Wörter oder Sätze, gesehen werden. Aber Menschen entwickeln auch visuelle Darstellungen nonverbaler Codierung für nichtsprachlichen Formen, die als Alltagsgegenstände oder Szenen wahrgenommen werden. Ebenso entwickeln Menschen auditive Darstellungen im verbalen Code für Spracheinheiten, die als Phoneme und deren Kombinationen gehört werden und auditive Darstellungen im nonverbalen Code für nichtsprachlichen Umweltgeräusche. Desgleichen werden haptische (d.h. kinästhetische oder taktile) Vertretungen im verbalen Code für sprachliche Bewegungsakte und haptische Darstellungen in der nonverbalen Codierung für das aktive „Fühlen“ von Objekten, Texturen und Bewegungen entwickelt. Chemische Sinnesmodalitäten (wie Geruch und Geschmack) werden bspw. nicht durch Sprache repräsentiert, aber es gibt für sie zum Teil nonverbal Darstellungen (z.B. der Geruch und Geschmack eines frisch gebackenen Kuchens bei einem Geburtstag). Bilder in dieser Modalität (Geschmack oder Geruch) sind in der Regel für die meisten Menschen weniger lebhaft. Dazu müssen Affekte wie emotionale Gefühle und Reaktionen hinzugefügt werden. Diese sind per definitionem nonverbal, obwohl es viele Namen für emotionale Zustände gibt. Des Weiteren existieren auch Bilder für solche Zustände, welche eine wichtige Komponente der Bedeutung darstellen. Diese verarbeitenden und speichernden Subsysteme scheinen unabhängig voneinander und spezialisiert in bestimmten manchmal multiplen Bereichen des Gehirns zu sein (vgl. Sadoski/Paivio 2004, S. 3f)⁵⁷.

Zur überschaubareren Darstellung entwickelten Sadoski/Paivio folgende Übersicht:

Tabelle 4 – Orthogonal Relationship between Mental Codes and Sense Modalities
(Sadoski/Paivio 2004, S. 5)

⁵⁷ Die Übersetzung wie auch die folgenden wurde vom Autor selbst verfasst.

Mental Codes

Sense Modality [Sinnesmodalität]	Verbal	Nonverbal
Visual [visuell]	Visual language (writing) [sichtbare Sprache (geschriebenes)]	Visual objects [sichtbare Dinge]
Auditory [auditiv]	Auditory language (speech) [hörbare Sprache (gesprochenes)]	Environmental sounds [Umweltgeräusche]
Haptic [haptisch]	Braille, handwriting [Braille-Schrift, Handgeschriebenes]	"Feel" of objects [„Fühlen“ von Dingen]
Gustatory [gustatorisch]	/	Taste memories [geschmackliche Erinnerungen]
Olfactory [olfaktorisch]	/	Smell memories [geruchliche Erinnerungen]

Empirische Belege für ihre Annahmen kommen vor allem aus der Neuropsychologie. So führen sie an, dass bspw. einige Personen mit Alexie⁵⁸ den Ausdruck Baseballschläger nicht lesen können, aber ihn erkennen, wenn er gesprochen wird und können ihn sogar schreiben, was laut ihnen den Nachweis von unabhängigen, modalitätsspezifischen Darstellungen innerhalb des verbalen Codes erbringt. Einige Personen mit Anomie⁵⁹ können einen Baseballschläger erkennen, sind jedoch nicht in der Lage ihn zu benennen, was ebenfalls als Nachweis einer allgemeinen Unabhängigkeit zwischen nonverbalen und verbalen Codes dienen kann. Für weitere relevante neuropsychologische Erkenntnisse wird zudem auf Paivio (1986) oder Sadoski et al. (1991) sowie auf West et al. (1998) verwiesen (vgl. Sadoski/Paivio 2004, S. 6).

Gemäß dieser Theorie gelangen Text- (und Bild)informationen zunächst über die Augen und soweit möglich über die Ohren in das sensorische Gedächtnis, wo die Daten für sehr kurze Zeit

⁵⁸ Eine Art Wortblindheit, sogenannte zentrale Buchstabenblindheit

⁵⁹ Eine Art Wortfindungsstörung

präsent gehalten werden, um über den weiteren Verarbeitungsprozess zu entscheiden. Für relevant befundene auditive Informationen werden über den auditiv/verbalen Kanal an das Arbeitsgedächtnis weiter geleitet, relevante visuelle Informationen werden über den visuell/bildhaften Kanal das Arbeitsgedächtnis weitergeleitet. Für geschriebene Wörter wird allerdings die Spezifität angenommen, dass sie zunächst im visuell/bildlichen Kanal verarbeitet werden und dann in den auditiv/verbalen Kanal transferiert werden (vgl. Sadoski/Paivio 2004, S. 16; Niegemann et al. 2004, S. 188, 192).

Somit müssten für geschriebenen Text kognitive Ressourcen beider Kanäle akquiriert werden, um die Informationen adäquat zu verarbeiten. Das Arbeitsgedächtnis ist der zentrale Ort des Gedächtnissystems, an dem multimediale Verarbeitung stattfindet. Es hält die Information, getrennt im visuell/bildhaften und auditiv/verbalen Kanal, präsent und ermöglicht eine aktive Verarbeitung der neuen Informationen. Auch wenn demgemäß Informationen üblicherweise über einen Kanal das Arbeitsgedächtnis erreichen, ist es manchmal auch während der Informationsverarbeitung notwendig den Kanal zu wechseln – bspw. bei der Verarbeitung von geschriebenem Text. Solche Sprünge zwischen den Kanälen (→ Moduswechsel) erfordern hohe kognitive Ressourcen (vgl. Sadoski/Paivio 2004, S. 16; Niegemann et al. 2004, S. 188, 192). Zudem müsste dies die Konstruktion eines mentalen Modells voraus setzen.

Modellhaft darstellen lassen sich die Vorgänge in etwa wie folgt:

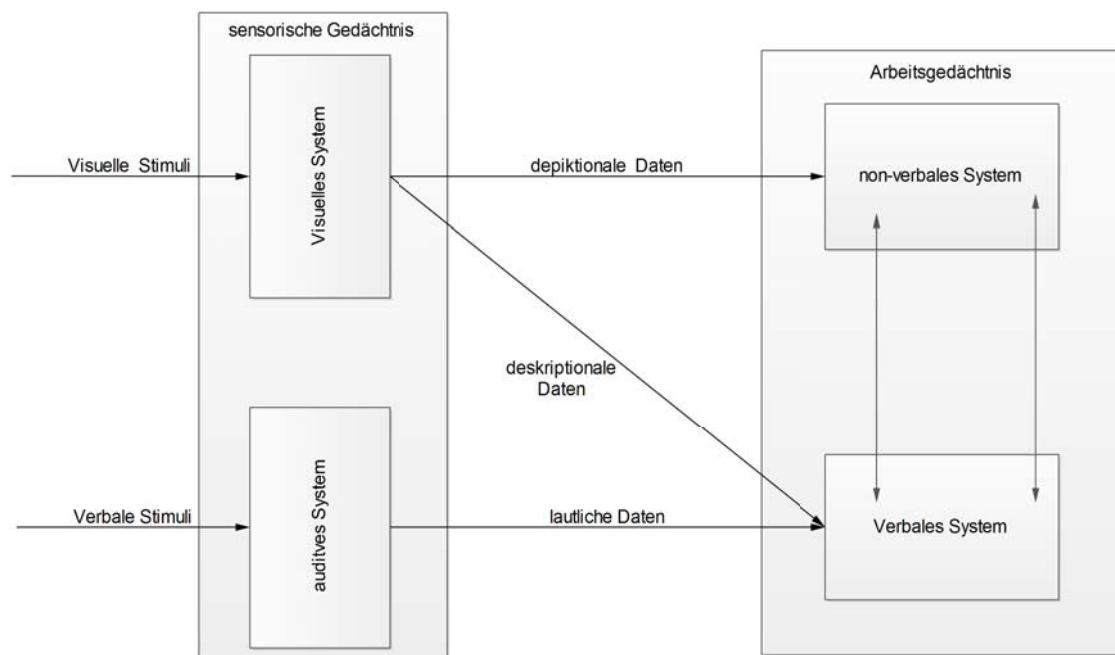


Abbildung 6 – Modell der DCT

In weiteren empirischen Studien⁶⁰ konnte jedoch weder eine Bestätigung der impliziten Annahmen zur Verwendung verbaler und piktorialer Information noch eine Bestätigung der expli-

⁶⁰ Bspw. Schnotz/Bannert (1999) – neuere Forschung fand bisher (Stand: 2015) nicht statt.

ziten Annahmen, wonach das Hinzufügen von Bildern zu einem Text eine generell lernunterstützende Funktion hat, gefunden werden. Stattdessen weisen einige Ergebnisse darauf hin, dass im Rahmen eines selbstgesteuerten Arbeitens (meist: Lernens) mit Texten das Angebot von depiktionalen Repräsentationen als weitere Informationsquelle nicht direkt zur Addition eines weiteren kognitiven Verarbeitungsprozesses führt. Es scheint eher so zu sein, dass Texte und Bilder vielmehr verschiedene Wege zur Konstruktion einer mentalen Repräsentation des dargestellten Konzepts bieten, welche einander teilweise substituieren, teilweise aber auch stimulieren können. Ferner weisen zusätzliche Resultate darauf hin, dass einfachere Bilder seitens der Subjekte eher eine oberflächlichere Verarbeitung evozieren, statt einer tieferen wie zuvor angenommen, bei der Text- und Bildverstehen einander teilweise substituieren, während diffizile Bilder konform zu den getroffenen Annahmen zu einer intensiveren Verarbeitung führen, bei der Textverstehen und Bildverstehen einander gegenseitig stimulieren (vgl. Schnotz/Bannert 1999).

Darüber hinaus scheint die duale Kodierungstheorie auch nicht in der Lage zu sein zu erklären, weshalb heterogene, aber informationsäquivalente Visualisierungsformen zu verschiedenen Wissensstrukturen mit differierenden Nutzungseigenschaften führen. Dagegen scheint das im Folgenden unten präsentierte integrative Modell des Text- und Bildverstehens eine eher geeignete Grundlage für die Erklärung solcher Effekte zu sein (vgl. Schnotz/Bannert 1999).

Schnotz/Bannert (1999; 2003), Schnotz/Dutke (2004) bzw. Schnotz (2005) entwickelten zur Erklärung und Analyse von Text-Bild-Interaktionen bei Verstehensprozessen das sogenannte „Integrative Modell des Text- und Bildverstehens (ITPC)“. In diesem Modell wird zentral zwischen Repräsentationen in Form von Propositionen und Repräsentationen in Form von mentalen Modellen unterschieden (vgl. Schnotz 2010, S. 67). Die Prämisse ist, dass Texte, Bilder und Diagramme auf unterschiedliche Weise zum Aufbau multipler mentaler Repräsentationen beitragen und dabei einander ergänzen können (vgl. Schnotz 2010, S. 68).

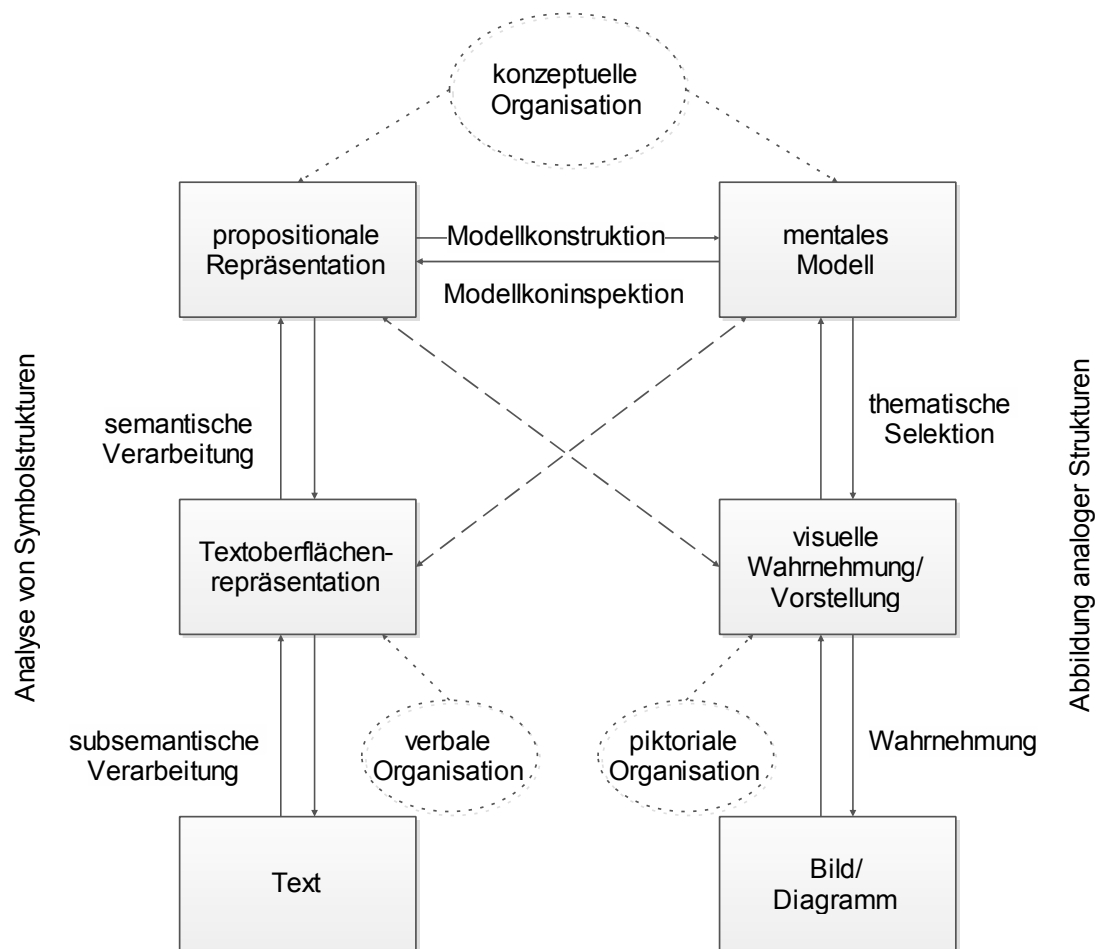


Abbildung 7 – Schematische Darstellung eines integrierten Modells des Text- und Bildverstehens (vgl. Schnotz/Bannert 1999; Schnotz/Dutke 2004)

Analog zu Sadoski/Paivio gehen Schnotz/Bannert von einem Modell mit zwei verschiedenen Zweigen der Verarbeitungswege aus:

Der deskriptionale Pfad (links) befasst sich mit der Analyse symbolischer Strukturen und besteht aus dem externen Text, der intern durch verbale Organisation erzeugten mentalen Repräsentation der Textoberflächenstruktur sowie der ebenfalls internen propositionalen Repräsentation und der semantischen Bedeutung, welche durch konzeptionelle Organisation erzeugt wurde. Die Interaktion zwischen diesen Repräsentationsschritten basiert auf Prozessen der Symbolverarbeitung. Der depiktionale Pfad (rechts) befasst sich mit der Analyse analoger Strukturen und besteht aus dem externen Bild, der auf piktorialer Organisation basierender internen visuellen Perzeption bzw. Vorstellung des Bildes sowie dem internen mentalen Modell des dargestellten Sachverhalts, welches wiederum aus einer konzeptionellen Organisation hervorgegangen ist. Die Interaktion zwischen diesen Repräsentationen basiert auf Prozessen der analogen Strukturabbildung (vgl. Schnotz/Bannert 1999, S. 3f).

Während des Lese- und Verstehensprozesses eines Texts konstruiert das Individuum demzufolge eine mentale Repräsentation der Textoberflächenstruktur, generiert weiterhin auf dieser Grundlage eine propositionale Repräsentation des semantischen Gehalts und formt anhand dieser sogenannten Textbasis schließlich ein mentales Modell des dargestellten Sachverhalts. Diese

Konstruktionsprozesse beruhen auf einer Interaktion von auf- und absteigenden Schemaaktivierungen, wodurch gleichzeitig Selektions- und Organisationsprozesse ablaufen können: Durch absteigende Aktivierung werden spezifische anforderungs- bzw. intentionsrelevante Daten selektiert und durch die Wechselbeziehung auf- und absteigender Aktivierungsprozesse kann sich jeweils eine bestimmte Schemakonfiguration formen, welche am besten zu den gegebenen Textinformationen passt und diese in einen kohärenten Komplex integriert. Verbale Organisationsprozesse verwerten Sprachinformationen gemäß syntaktischer und morphologischer Aspekte und generieren eine strukturierte mentale Repräsentation der Textoberfläche. Konzeptuelle Organisationsprozesse verwerten die Textoberflächenrepräsentation und generieren eine strukturierte propositionale Repräsentation und ein mentales Modell (vgl. Schnotz/Bannert 1999, S. 4).

Bei der Betrachtung und dem Verstehen einer depiktionalen Repräsentation erzeugt das Individuum durch Wahrnehmungsprozesse eine visuelle mentale Repräsentation des Bildes. Durch semantische Verarbeitungsprozesse werden weiterhin ein mentales Modell sowie eine propositionale Repräsentation des dargestellten Gegenstands aufgebaut. Bei der Bildverarbeitung werden durch absteigende Schemaaktivierung anforderungsrelevante Informationen selektiert und aufsteigend durch automatisierte visuelle Routinen enkodiert. Diese perzeptuelle Enkodierung beinhaltet auch die Identifikation und Diskrimination grafischer Elemente sowie Prozesse der visuellen Organisation. Die so gebildete mentale Repräsentation entspricht der visuellen Wahrnehmung des Bildes. Da Perzeption und Vorstellung auf den gleichen kognitiven Mechanismen und Prinzipien beruhen, kann die äquivalente Form der Repräsentation in anderen Kontexten auch als Vorstellung verstanden werden, indem die Generierung nicht über externe sensorische Informationen, sondern über interner Wissensstrukturen erfolgt. Damit eine Abbildung nicht nur perzeptiert, sondern auch rezipiert werden kann, müssen semantische Verarbeitungsprozesse ablaufen. Das Individuum konstruiert sodann ein mentales Modell des repräsentierten Gegenstands, indem wahrgenommene depiktionale Informationen auf mentale Entitäten und räumliche Verhältnisse auf semantische Relationen abgebildet werden. Das Bildverstehen kann demzufolge als ein Prozess der schemageleiteten analogen Strukturabbildung eines Systems von visuell-räumlichen Relationen auf ein System von semantischen Relationen verstanden werden. Dieser Abbildungsprozess kann in beiden Beziehungen stattfinden. Beim Verstehen realistischer Bilder wird auf vorhandene Schemata der alltäglichen Perzeption Bezug genommen, weshalb Weidenmann (1988) von einem ökologischen Bildverstehen spricht (siehe oben). Hingegen sind für die korrekte Verarbeitung logischer Bilder spezifische kognitive Schemata (Grafik-Schemata) notwendig, um an den visuell-räumlichen Konfigurationen Informationen über den dargestellten Gegenstand ablesen zu können (vgl. Schnotz/Bannert 1999, S. 4f).

Empirische Studien konnten bereits Hinweise darauf liefern, dass die Konstruktion mentaler Modelle das Aufnahmevermögen des imaginalen Arbeitsgedächtnisses auslastet. Jedoch sind mentale Modelle, da es sich um eine sensorisch unspezifische Form der depiktionalen Repräsentation handelt, abstrakter als visuelle Vorstellungen. Überdies unterscheiden sich mentale Modelle von visuellen Vorstellungen hinsichtlich ihrer Informationsdichte. Erstens basiert die

Konstruktion eines mentalen Modells beim Bildverstehen auf einem anspruchsbabhängigen Selektionsprozess. Da im Allgemeinen die Verarbeitungsprozesse an die antizipierte Aufgabenstellung sowie Anforderungsstruktur angepasst werden, werden nur die Komponenten der depiktionalen Konfiguration in den Prozess der Strukturabbildung einbezogen, die für die Verarbeitung der antizipierten Anforderungen relevant scheinen. Zweitens kann davon ausgegangen werden, dass mentale Modelle durch die Benutzung von Weltwissen elaboriert werden und so auch Informationen enthalten können, welche im Originalbild nicht vorhanden sind. Der Nutzen eines mentalen Modells ist, dass theoretisch durch Prozesse der Modellinspektion neue Informationen abgelesen werden können, die in keiner der ursprünglichen Repräsentation explizit vorhanden waren. Diese gewonnenen Ableseergebnisse werden anschließend in Form von Propositionen enkodiert und expliziert. Demzufolge findet auch zwischen einer propositionalen Repräsentation und einem mentalen Modell eine unablässige Interaktion von Modellkonstruktions- und Modellinspektionsprozessen statt (vgl. Schnotz/Bannert 1999, S. 5).

Beim Textverstehen wird davon ausgegangen, dass der Ursprung und die Basis der Wechselbeziehung eine propositionale Repräsentation ist, anhand derer das mentale Modell konstruiert wird. An diesen werden dann wieder neue Informationen abgelesen und können der propositionalen Repräsentation hinzugefügt werden. Beim Bildverstehen dagegen ist der Anfangspunkt der Wechselbeziehung ein mentales Modell, an welchem durch Ableseprozesse neue Propositionen gewonnen werden. Demzufolge besteht zwischen den Arten der externen Darstellung und den Arten der internen mentalen Repräsentation keine Eins-zu-eins-Relation. Respektive kann ein Text als externe deskriptionale Repräsentation intern sowohl zu einer deskriptionalen als auch zu einer depiktionalen Repräsentation führen – vice versa kann ein Bild als externe depiktionalen Repräsentation intern sowohl zu einer depiktionalen als auch zu einer deskriptionalen Repräsentation führen (vgl. Schnotz/Bannert 1999, S. 5).

Rein methodisch-formal betrachtet wird offenkundig, dass die Konstruktion einer propositionalen Repräsentation und eines mentalen Modells ebenfalls als eine duale Kodierung charakterisiert werden kann. Die an dieser Stelle explizierten Modellannahmen unterscheiden sich jedoch entscheidend von der traditionellen dualen Kodierungstheorie. Zunächst wird hier nicht nur für die Verarbeitung von depiktionalen Strukturen, sondern auch für die Verarbeitung von Sprache eine duale Kodierung postuliert. Überdies wird die Konstruktion eines mentalen Modells nicht nur als eine Ergänzung einer weiteren Kodierung verstanden, welche einen quantitativen Nutzen gegenüber einer einfachen Kodierung erbringt, sondern als wesentlich wird demgegenüber angesehen, dass propositionale Repräsentationen und mentale Modelle auf unterschiedlichen, einander ergänzenden Repräsentationsprinzipien beruhen (vgl. Schnotz/Bannert 1999, S. 5).

Auf Grund der kurz dargestellten Vor- und Nachteile der jeweiligen Theorie sowie ihrer empirischen Belege oder Widersprüche und weiterer Implikationen scheint es an dieser Stelle angebracht der Modellierung von Schnotz/Bannert zu folgen und das integrative Modell des Text- und Bildverstehens als bisher plausibelste Erklärung der Prozesse der gemeinsamen Text-/Bild-

verarbeitung anzusehen. Gemeinsam und zentral ist die Annahme, dass ein Wechsel des Repräsentationsformats eine kognitive Verarbeitung, wahrscheinlich auf der Ebene des mentalen Modells, erfordert.

Fraglich bleibt noch, welche Probleme beim Ablauf dieser Prozesse auftreten und welche Faktoren sie vereinfachen können. Dem widmet sich das nächste Kapitel.

3.1.5. Aspekte der Textverständlichkeitsforschung und ihre Übertragbarkeit auf depiktionale Repräsentationen

Es haben sich verschiedene Systeme entwickelt, um schwierigkeitsbestimmende Parameter von Texten zu ermitteln. Dabei gilt zu beachten, dass sich das im Folgenden dargestellte nahezu ausschließlich auf deskriptionale Repräsentationen bezieht – Ergebnisse zur Anforderungen an die Verarbeitung depiktionaler Repräsentationen liegen bisher nicht vor, sind aber auf Grund der in Kapitel 2.2 gezeigten Erkenntnisse dringend notwendig. Ob eine Äquivalenz der Parameter für depiktionale und Hybridformen angenommen werden kann, wird an der jeweiligen Stelle in die Reflexion einbezogen.

Ziel des Kapitels ist es mögliche Parameter herauszuarbeiten, die zumindest teilweise die Kontrolle über die Schwierigkeit von Leseaufgaben ermöglichen. Eine konkrete Prüfung der Anwendbarkeit sowie ggf. Adaption für schwierigkeitsbestimmende Itemparameter erfolgt bei der Konstruktion und Analyse von Items in Kapitel 5.1.

Zunächst ist kurz Prägnanz zu schaffen und die Textverständlichkeitsforschung von der allgemeineren Lesbarkeitsforschung abzugrenzen.

Die Lesbarkeitsforschung entstanden in den 20er Jahren des 20. Jahrhunderts im anglo-amerikanischen Sprachraum definiert Lesbarkeit wie folgt:

„Ein Text sollte lesbar genannt werden, wenn er auf der Ebene der lokalen Wahrnehmungseinheiten (an der textuellen Oberfläche) der Wahrnehmung des Lesers keine oder nur geringe Schwierigkeiten bereitet.“ (Sauer 1995, S. 163).

Diese Definition macht deutlich, dass Lesbarkeit im Gegensatz zu Verstehbarkeit einerseits auf die Textoberfläche, d.h. die Außenseite des Textes, bzw. andererseits auf die Wahrnehmung des Lesers bezogen ist. Damit ist ein nicht entzifferbarer Text ein Text, der unabhängig von den jeweiligen Kompetenzen des Lesers nicht decodiert werden kann, bspw. weil die Schriftart zu klein gewählt wurde oder weil Verschmutzungen zu Beeinträchtigungen führen (vgl. Sauer 1995, S. 163). Zumindest diese Kriterien dürften allgemein akzeptiert werden und daher auch für depiktionale Texte angenommen werden.

Die Lesbarkeitsforschung liefert zwar durch verschiedene Berechnungen von Lesbarkeitsformeln, welche auf quantitativ auswertbaren und objektiv feststellbaren Textmerkmalen, die sta-

tistisch zueinander in Verbindung gesetzt werden⁶¹, basieren, interimistisch plausible Ergebnisse, jedoch ist die Reliabilität solcher Formeln meist nicht überprüft worden (vgl. Groeben 1982, S. 177). Besonders problematisch ist aber, dass die Validitätstests ebendieser Formeln wenig gute Ergebnisse hervorbrachten (vgl. Groeben 1982, S. 178f).

Da bei diesen originär für die englische Sprache entwickelten Formeln lediglich auf einfache Oberflächenmerkmale wie Lexik oder Syntax eingegangen wird, wird die semantisch-pragmatische Konstruktion eines Textes, das meint bspw. die Kohärenzbildung oder die grammatische Korrektheit von Sätzen, nicht bei der Beurteilung der Lesbarkeit berücksichtigt. Dies ist an dieser Herangehensweise als wesentliche Kritik zu nennen. Zusätzlich kommt Willenberg zu der Feststellung, dass die für den englischen Sprachraum entwickelten Lesbarkeitsformeln u.a. auf Grund der komplexeren Sätze und anderen Wort- bzw. Silbenstrukturen sich nicht für eine Übertragung ins Deutsche eignen (vgl. Willenberg 2005, S. 95).

Wesentliche Erkenntnisse, die bei der Betrachtung schwierigkeitsbestimmender Parameter nicht unberücksichtigt bleiben sollten, sind dennoch, dass es die Lesbarkeit eines Textes vereinfacht, wenn auf der Wortebene geläufige und gebräuchliche Wörter, d.h. Wörter...

- ... die früh in der Sprachentwicklung gelernt werden
- ... die kurz (gemäß Buchstaben und/oder Silbenlänge) sind
- ... die nicht in die Klasse der Fremdworte bzw. Termini technici fallen
- ... die für Schriftsprache (nicht mündliche Sprache) gebräuchlich sind
- ... die eine eingeführte, gebräuchliche Bedeutung besitzen
- ... die konkret und anschaulich, statt abstrakt sind

verwendet werden (vgl. Groeben 1982, S. 185).

Auf der Satzebene kann vor allem die Komplexität durch die Verwendung möglichst kurzer und grammatikalisch einfacher Sätze, d.h. ohne komplexe und komplizierte Satzschachtelungen, insbesondere Hypotaxen⁶², reduziert werden (vgl. Groeben 1982, S. 185).

Diese textuellen Oberflächenmerkmale deskriptionaler Texte sind allerdings bereits nicht mehr ohne weiteres auf depiktionale Strukturen übertragbar, da in depiktionalen Repräsentationen klar definierte Referenzen zu Buchstaben, Wörtern und Sätzen fehlen. Eine Möglichkeit der Betrachtung wäre es sicherlich, dass eingeführte, gebräuchliche und häufig verwendete Symbole einfacher zu verstehen sind, genauso wie eindeutige Kontraste.

⁶¹ Beispiele: Dale-Chall-Formel, Reading Ease und Human Interest nach Flesch, Farr-Jenkins-Peterson Formel oder Dickes-und-Steiwer-Formel sowie die Formel von Reiners.

⁶² untergeordnete Nebensätze

Dennoch kann Lesbarkeit bezogen auf die Oberfläche betrachtet sicherlich als eine notwendige Voraussetzung für Verständlichkeit angenommen werden – unabhängig vom Repräsentationsmodus.

Verständlichkeit oder Verstehbarkeit eines Textes im nächsten Schritt kann definiert werden als leserseitige kognitive Verarbeitbarkeit lokaler Einheiten (vgl. Sauer 1995, S. 163). Sauer führt dazu weiter aus, dass die Verständlichkeit somit direkt auf das Vorwissen des jeweiligen Lesers Bezug nimmt und sich dabei im Spannungsfeld von überfordernder Neurekonstruktion und inhaltlicher Anspruchslosigkeit bewegt (vgl. Sauer 1995, S. 164).

Teigeler hielt bereits in den 60er Jahren verschiedene Merkmale fest, die er verschiedenen Klassen zuordnet, die einzeln durch empirische Analysen gestützt werden. Schwierigkeitsbestimmende Parameter von Texten wären dem entsprechend (Teigeler 1968, S. 40–53):

- Semantische Aspekte (Vokabular)
- Abstrakte Wörter
- Fachausdrücke
- Fremdwörter
- Das Verhältnis von leichten zu schweren Wörtern
- Semantische Aspekte (Textzusammenhang)
- Induktives statt deduktives Vorgehen
- Übergänge
- Syntaktische Aspekte
- Satzlänge
- Satztiefe⁶³
- Satzart⁶⁴

Diese sind zwar textueller Natur und empirisch überprüft, jedoch nicht in Kombination miteinander, so dass unklar ist welches dieser Merkmale den stärksten Einfluss auf das Textverstehen hat und welche eher randständige Merkmale sind. Die Anwendbarkeit bezüglich depiktionaler Repräsentationen ist nicht gegeben.

Ausgehend von der Frage, welche Prozesse rezipientenseitig ablaufen müssen, damit ein Text verstanden wird und was diese Prozesse behindert, liefert Wolff (1985), einige plausible Merkmale bezüglich der Schwierigkeit.

⁶³ Anzahl der grammatischen Fakten, die ein Leser verarbeiten muss

⁶⁴ Passiv statt aktiv Konstruktion; Substantivierung statt verbale Formulierung; Nominalisierungen;

- Die Schwierigkeit sei erstens abhängig von den Möglichkeiten, die dem Rezipienten zur Verfügung stehen oder zur Verfügung gestellt werden, um eine Eingangshypothese über den Text zu formulieren (Konstruktion eines allgemeinen Modells des zu erwartenden Textes, Aktivierung von Vorwissen,...) (Wolff 1985, S. 215) → Leseraktivität
- zweitens könne die Textschwierigkeit auf der Basis der dem Text zugrundeliegenden Konzepte und Propositionen beurteilt werden (Wolff 1985, S. 216) → Textmerkmal
- drittens könne die Schwierigkeit beurteilt werden nach dem Textrezipienten zur Verfügung stehenden Textsortenwissen (Wolff 1985, S. 216) → Lesermerkmal (Struktur)
- viertens könne die Schwierigkeit beurteilt werden auf Basis der im Text angelegten Kohärenzdichte und abschließend (Wolff 1985, S. 216) → Textmerkmal
- fünftens auf der Basis des dem Rezipienten zur Verfügung stehendem Weltwissen und der Möglichkeit der Anwendbarkeit dieses Wissens während der Textverarbeitung (Wolff 1985, S. 216) → Lesermerkmal (strukturell)

Diese Merkmale entwickelte er für die Frage nach der Schwierigkeit eines fremdsprachlichen Textes, dennoch kann angenommen werden, dass einen muttersprachlichen Text ähnliche Parameter anzunehmen sind. Bezüglich der Operationalisierbarkeit dieser Kriterien bleibt allerdings nur sein Eingeständnis, dass dies nicht möglich sei (vgl. Wolff 1985, S. 220). Eine Erklärung dafür kann darin liegen, dass er unterschiedliche Referenzsysteme verwendet: Die Möglichkeiten, die dem Rezipienten zur Verfügung stehen, um eine Eingangshypothese über den Text zu formulieren beziehen sich auf die konkrete Situation, die dem Text zugrundeliegenden Konzepte und Propositionen sowie die Kohärenzdichte sind textseitige Merkmale, das zur Verfügung stehende Textsortenwissen und das Weltwissen sowie die jeweilige Anwendbarkeit sind leserseitige Faktoren.

Interessant am Stand der Textverständlichkeitsforschung ist – wie bereits an Hand der frühen Arbeiten erkennbar ist –, dass die Verständlichkeit eines Textes expliziert wird als mediatives Konstrukt, welches die Verbindung zwischen den Merkmalen der Textstruktur und den Behaltenskriterien herstellt (vgl. Christmann/Groeben 1996, S. 173). Hierin wird deutlich wie sehr in der angewandten Linguistik das Behalten eines Textes als zentrale Funktion und quasi ausschließlich die Leserintention des Lernens fokussiert wird. Fraglich ist, ob dies für alle in Kapitel 3.2.2 beschriebenen Intentionen angenommen werden kann – dies darf wahrscheinlich bezweifelt werden.

Dennoch werden auf Grund des Fehlens anderer Theorien kurz die beiden grundlegenden Wege zur Feststellung relevanter Textmerkmale und ihrer Ergebnisse gemäß der Textverständlichkeitsforschung vorgestellt, die sich allerdings wie bereits erwähnt exklusiv auf deskriptionale Texte beziehen.

Der erste Weg ist der von Langer, Schulz von Thun/Tausch (1974) als empirisch-deduktive beschrieben und bereits früh eingeschlagen wurde. Es wurden unterschiedlich schwierige Texte von Experten auf einer siebenstufigen bipolaren Skala bezüglich ihrer Meinung nach relevanter Schwierigkeitsmerkmale auf Text- wie auf Stilebene eingeschätzt und anschließend eine Faktoranalyse der Ratingdaten durchgeführt. Diese ergab folgenden vier Dimensionen:

1. Sprachliche Einfachheit (Darstellung, Satzlänge, Geläufigkeit der Wörter, Erklärungen,...)
2. Gliederung bzw. Ordnung (Gliederung, Folgerichtigkeit, Übersichtlichkeit, roter Faden,...)
3. Kürze bzw. Prägnanz (Wesentlichkeit, Drängung, Lernzielbezug,...) und
4. zusätzliche Stimulanz (Anregung, Abwechslung, Persönlichkeit,...) (vgl. Christmann/Groeben 1996, S. 174).

Anschließend wurde die Ausprägung eines Textes auf jeder dieser Dimensionen mittels einer fünfstufigen bipolaren Rating-Skala eingeschätzt. Die daraus resultierenden vier Werte geben das quantitative Ausmaß der Verständlichkeit an. Daraufhin optimierte Texte wurden von Probanden signifikant besser behalten als die jeweiligen Originaltexte. Damit kann laut Christmann/Groeben den vier Dimensionen zumindest ein praktischer Erklärungswert zugeschrieben werden. Probleme dieses Ansatzes sind die verhältnismäßige Theoriearmut⁶⁵ und der alleinige Einsatz von Experten-Ratings zur Bestimmung der Parameter. Weiterhin seien die ermittelten Merkmale zu wenig explizit, als dass sie als konkrete Konstruktionsanweisungen für die Produktion allgemein als verständlich geltender Texte ableitbar wären (vgl. Christmann/Groeben 1996, S. 174). Fragwürdig sind auch die sprachlich unscharfen Definitionen der Kategorien, welche wahrscheinlich zu intuitiven Raterurteilen geführt haben. Daneben sind klare dimensionale Strukturen nicht in wünschenswerter Weise gegeben, da bspw. Kürze und Textverstehen nicht in jedem Fall in einem eindeutig positiven Zusammenhang stehen (vgl. Willenberg 2005, S. 95). Zudem wurde wiederum versucht das Konstrukt Textverstehen über eine Behaltensleistung zu ermitteln.

Der andere Weg ist der von Groeben beschrittene Weg der theoretisch-deduktiven Bestimmung der Kategorien auf Grundlage sprachpsychologischer Theorien, lerntheoretischen und motivationspsychologischen Ansätzen zur Textrezeption. Auch dieser Weg führt zu vier Verständlichkeitsdimensionen deren Textmerkmale

- stilistische Einfachheit (kurze Satzteile, aktive Verben, keine Nominalisierungen,...),
- semantische Redundanz (bedachte Wiederholung, Weitschweifigkeit), kognitive Strukturierung (Hervorhebung, Zusammenfassung, Beispielgebung,...) und

⁶⁵ So dann wird fraglich, wieso er von den Autoren als deduktiver Weg bezeichnet wird

- kognitive Konflikterzeugung (Neuheit und Überraschung, inkongruente Konzepte, Fragen,...) sind.

Diese postulierten Dimensionen wurden mittels eines varianzanalytischen Verfahrens mit 18 Texten, die inhaltlich identisch waren, allerdings hinsichtlich der Faktoren variierten, untersucht. Bei dieser Messung wurden Verständlichkeit, Behalten und Interesse als abhängige Variablen erhoben. Dabei bietet sich das informationstheoretische Modell für die Betrachtung an, da darin

„die Mediationsfunktion des Verständlichkeitskonstrukts charakterisiert [werden kann]; sie besteht nicht nur in einer Integration der verschiedenen Dimensionen der Textstruktur in den einheitlichen Gradmesser subjektiver Information, der über den Rezeptionsprozess die erklärende Brücke zu den Lernkriterien schlägt. Darüber hinaus gestattet der Rückbezug auf die individuellen Erwartungsstrukturen die Einbeziehung der Zeitdimension und des durch sie bedingten möglichen inhaltlichen Wandels der Interpretation von Gegenstand und Rezipient.“ (Groeben 1978, S. 76)

Das hier relevante Konstrukt der Verständlichkeit eines Textes wird sodann über einen Ratevorgang⁶⁶ auf Ebene der Worteinheit mit einem Zeitlimit (30 Sekunden) operationalisiert, um so die subjektiven Erwartungsstrukturen eines Rezipienten einbeziehen zu können, statt nur auf quantitative Maße der syntaktischen Textstruktur zu referieren (vgl. Groeben 1978, S. 79). Die Ergebnisse der durchgeführten Varianzanalyse lassen die Interpretation zu, dass für die Verständlichkeit eines Textes von den obigen Dimensionen der Faktor der inhaltlichen Strukturierung der Ausschlag gebende sei (Varianzaufklärung: 86%)⁶⁷ (vgl. Christmann/Groeben 1996, S. 174). Willenberg kritisiert an diesem Ansatz einerseits die geringe empirische Basis und, dass nicht eindeutig bestimmbar war, dass die angenommenen Subdimensionen konsequent negative Auswirkungen auf die Verständlichkeit haben (vgl. Willenberg 2005, S. 95).

Auch wenn beide Wege im Vergleich aufzeigen, dass sie bezüglich der Dimensionen zu ähnlichen Erkenntnissen kommen, wird jedoch die bereits angesprochene Engführung auf epistemische Kontexte durch die Fokussierung auf Lernziele, Motivation oder eine Behaltensleistung deutlich. Es lässt sich hier fast einzig eine optimal gelungene didaktische Aufbereitung als Zielkategorie des verständlichen Textes erwarten. Damit ist fraglich, ob diese Merkmale auch für nicht didaktisierte Texte gelten und damit allgemeinere Anforderungsstrukturen abbilden können.

⁶⁶ Zur Präzisierung: gemeint ist ein Vorgang in dem etwas erraten wird.

⁶⁷ (Christmann und Groeben 1996) führen weiter aus, dass daneben sprachliche Einfachheit 3,5% der Varianz signifikant aufklärte; die semantische Redundanz sich nur in Kombination mit der sprachlichen Einfachheit auf das Verstehen auswirkte und für das Behalten nur der Faktor der inhaltlichen Strukturierung relevant war (vgl. Christmann/Groeben 1996, S. 175).

Weiterhin hält Kirscher (1998) fest, dass bestimmte sprachliche Mittel in jedem Fall besondere Schwierigkeiten bereiten. Dazu gehören Konjunktionen und Subjunktionen, Komposita, Fachtermini, Pro-Formen⁶⁸, hypotaktischer Satzbau, Funktionsverbgefüge sowie Passivkonstruktionen. All dies fasst sie unter dem Begriff der Textdichte zusammen (vgl. Kirscher 1998). Auch wenn diese Aspekte theoretisch plausibel zu sein scheinen, wurden bisher keine diesbezüglichen empirischen Untersuchungen durchgeführt, so dass nicht ermittelt ist, ob die angenommenen Parameter tatsächlich schwierigkeitsbestimmend sind und falls ja, in welchem Umfang. Eine Übertragung auf depiktionale Repräsentationsformate ist weiterhin auch kaum möglich.

Ein weiterer Ansatz zentraler Kategorien zur Bestimmung der Textschwierigkeit wurde von Willenberg (2005) vorgelegt. Er kam zu sechs Aspekten, die für verschiedene Texte anwendbar und von denen fünf auch auf Sachtexte übertragbar seien. Diese wären im Einzelnen (vgl. Willenberg 2005, S. 95):

1. Satzlänge im so genannten „Drei-Sekunden-Fenster“
2. Wortschatz auf vier Ebenen (Basiswörter, Konkreta, Abstrakta, Fachwörter)
3. Vertextung mit Junktoren innerhalb von Sätzen und in der Verbindung von Sätzen
4. Redundanz basierend auf Schlüsselwörtern
5. Verlebendigung (Personen, Handlungen und/oder Emotionen innerhalb der Texte)
6. Literarisierende Merkmale (Metaphorik, Perspektivik, uneigentliche Redeweisen wie Ironie, Komik, Übertreibungen und Mehrdeutigkeiten)

Ziel war es, nach Möglichkeit genaue, sprachlich einfache Kategorien mit kohärenter Begründung zu entwickeln, die ebenfalls für empirische Verfahren nutzbar sind (vgl. Willenberg 2005, S. 95).

Die gefundenen Merkmale begründet er theoretisch durch Erkenntnisse aus Neuropsychologie und Leseforschung. So kann er festhalten, dass die Herstellung von Bedeutung auf dem Finden eines genauen Knotens im Wissensnetz und im nächsten Schritt auf der Aktivierung seiner Merkmale und seiner Vernetzungen beruht. Daneben müssen Bedeutungs-Konkurrenten frühzeitig ausgeschlossen werden. Nach seinen Erläuterungen unterstützen seine Kategorien diese Prozesse⁶⁹ (vgl. Willenberg 2005, S. 99). Auch hier wiederum wird offenkundig, dass eine Übertragung auf nicht deskriptionale Formate kaum möglich ist. Dennoch ist dies einer der Ansätze, die tatsächlich auf der Verstehensebene operieren und nicht nur das Behalten der propositionalen Struktur des Textes erfassen.

⁶⁸ Sprachliches Mittel der Wiederaufnahme, ohne die Verwendung des Begriffs selbst (Pronomen, Synonyme, Oberbegriffe, Präpositionalobjekte)

⁶⁹ Neben diesen genannten Aspekten wird in der Psycholinguistik auch Art und Umfang der Inferenzen, die ein Leser eines Textes ziehen muss, um ihn zu verstehen, als wichtiges Element der Verständlichkeit genannt (vgl. Krings 1996, S. 39f).

Einer der weiteren wenigen Ansätze, der wie Willenberg bzw. DESI auch andere kognitive Verarbeitungsprozesse als das Memorieren und deren Überprüfung durch Wiedergabe berücksichtigt, ist das Situationsmodell oder Mentale Modell nach Kintsch (1994), welches von Rickheit (1995) als aufgabenorientierter Ansatz klassifiziert wird. Dieser orientiert sich an sogenannten Real-Life-Situationen und versucht eine theoriegeleitete und methodenbewusst arbeitende praxisnahe Forschung zu ermöglichen. Ausgehend von der Beobachtung, dass Personen zwar in der Lage sein können einen Text zu reproduzieren, jedoch nicht in der Lage zu sein scheinen bestimmte Informationen für andere Zwecke zu nutzen, stellt sich die Frage nach dem Einfluss des Vorwissens auf Leseprozesse (vgl. Rickheit 1995, S. 24f). In verschiedenen Studien (Müsseler/Rickheit/Strohner 1985; Rickheit/Strohner/Müsseler 1987; Grabowski 1991) konnte bei einer Reanalyse die postulierte Interaktion des Weltwissens mit dem Textwissen bestätigt werden, jedoch in der Form, dass es einen so starken Effekt hat, dass es fast alle möglichen Schwierigkeitsbestimmenden Parameter überdeckt. Damit konnte vor allem festgehalten werden, dass die Verständlichkeit eines Textes nicht nur vom Text selbst abhängt, sondern auch oder vor allem von den kognitiven Prozessen, mit denen er verarbeitet wird, sowie von situativen und aufgabenspezifischen Rahmenbedingungen (vgl. Rickheit 1995, S. 25). Daher kommt Rickheit zu dem Schluss, dass bei der anschließenden Erforschung von Verständlichkeit von Texten neben sprachlichen und kognitiven Prozessen zunächst die Eingebundenheit in bestimmte aufgabenorientierte Szenarien stärker zu berücksichtigen sei und insbesondere auch andere als linear-sequentielle Texte einbezogen werden müssen (vgl. Rickheit 1995, S. 26).

Auch wenn die vorgestellten Ergebnisse den Ausgangspunkt und die Absicht der vorliegenden Arbeit stärken, bleibt doch die Frage offen, was macht einen Text zu verstehen so schwer, dass manche daran scheitern. Aus den bisherigen Ausführungen wird zumindest deutlich, dass die Frage, was die Schwierigkeit eines Textes und damit im weiteren Sinne einer Leseaufgabe ausmacht, weniger trivial ist als möglicherweise im Vorhinein angenommen wird. Daher kann sie auch an dieser Stelle nicht abschließend beantwortet werden, dennoch seien hier einige Hinweise aus der vorhergehenden Analyse zusammengetragen. Zunächst wird festgehalten und im Folgenden als gegeben betrachtet, dass ein Einfluss folgender textseitiger Aspekte⁷⁰ wahrscheinlich ist;

- Oberflächenstruktur, wie der Länge des Textes/Anzahl der Wörter
- Semantische Aspekte (Vokabular), wie die Wortwahl, sprachliche Einfachheit
- Syntaktische Aspekte, wie die Satzlänge und der Satzbau
- Kognitive Strukturierung, d.h. Gliederung, Ordnung, semantische Aspekte des Textzusammenhangs
- (Anforderungen an die) Inferenzbildung⁷¹, Semantische Redundanz und Kohärenzdichte bzw. Kohärenzlücken

⁷⁰ Dies ist nicht unabhängig von Leser- und Situationsmerkmalen abschließend beantwortbar (s.u.).

⁷¹ Anforderungen siehe unten

- Prägnanz
- Stimulanz oder Verlebendigung
- Literarisierende Merkmale
- Anreize zur Bildung einer Eingangshypothese
- Inhaltlicher Anspruch des Textes

Auf Seiten des Lesers wird verwiesen auf⁷²:

- Vor- bzw. Weltwissensbezug
- Textsortenwissen

Es lassen sich jedoch keine eindeutigen Zusammenhänge oder Gewichtungen aus der bisherigen Forschung ableiten, die eine Schwerpunktsetzung erlauben.

In wie weit alle diese Aspekte bei der später vorzustellenden Itemkonstruktion⁷³ berücksichtigt oder operationalisierbar gemacht werden können und welche weiteren Einflussgrößen einzu-beziehen sind wird in Kapitel 4.1.3 dargestellt.

3.1.6. Lesekompetenzmessung – Stand, Testverfahren und Pas-sung auf den beruflichen Kontext

Nachdem in Kapitel 3 und 0 ein Modell zur möglichen Erklärung von Leseprozessen dargestellt wurde sowie verschiedene Ansätze zur Synthese ebendieser vorgestellt wurden (vgl. Kapitel 0), müssen diese Prozesse noch operationalisiert werden, um vergleichenden Analysen zu ermög-lichen. Ziel dieses Kapitels ist es zu ermitteln, ob ein Verfahren zur Messung von Lesekompetenz – wie sie im beruflichen Bereich dominant zu sein scheint⁷⁴ – bereits existiert bzw. ob etablierte Verfahren für diesen Kontext adaptierbar scheinen.

Nach Bhola (1995) liegt das Problem bei der Lesekompetenzmessung schon darin, dass zum Teil ideologische und konzeptionelle Schlachten um die Seele des Lesens wüten, während es für Lese-Fachleute die praktische Notwendigkeit gibt für praktikable Definitionen bezüglich Al-phabetisierung und Lesefähigkeit zu sorgen, die auch verwendet werden können, um Nicht-Literaten von Literaten zu trennen und um weiterhin in der Lage zu sein, zwischen den ver-schiedenen Ebenen der Leistung von Einzelpersonen und Gruppen zu differenzieren (vgl. Bhola 1995, S. 7). Bhola resümiert “Nether the definition of functional literacy as skills of navigating in one’s culture, nor the definition of generalized functional literacy have proved to be useful.”⁷⁵

⁷² Weiterhin lassen sich Motivation, Handlungsabsicht, Lesestrategiewissen etc. als einflussfaktoren annehmen, dazu ist die For-schungslage jedoch zum Zeitpunkt des Verfassens noch prekärer.

⁷³ Siehe Kap. 4.1 Entwicklung des Testinstruments – S. 179

⁷⁴ Modelleinführung siehe Kap. 4.1.3 Theoretisches Kompetenzmodell – S. 187

⁷⁵ Weder die Definition der funktionalen Lesefähigkeit als Fähigkeiten zur Kursbestimmung in der eigenen Kultur, noch die Defi-nition der verallgemeinerten funktionale Lesefähigkeit hat sich als nützlich erwiesen.

(vgl. Bhola 1995, S. 9). Er hält diesbezüglich summierend fest, dass die Konzepte der Definitionen schwierig zu operationalisieren sind, insbesondere, da diese mit den Bedingtheiten der Sprache, dem Inhalt und dem Schwierigkeitsgrad des Textes, beruflichen Erfordernissen, sozialen Rollen und kulturellen Anforderungen verknüpft sind (vgl. Bhola 1995, S. 9).

Bei der Erstellung verschiedener Diagnoseinstrumente stand zu Beginn der Betrachtungen das zeitige Erkennen von Defiziten bezüglich der Lesefähigkeit im Fokus. Etliche Längsschnittstudien konnten Hinweise auf den überdauernden Einfluss früher Leselernprozesse liefern. So konnte beispielsweise von Klicpera/Gasteiger-Klicpera (1993) gezeigt werden, dass bei der Entwicklung des Lesetempos von der zweiten bis zur achten Schulklasse die Leistungsdifferenz zwischen Schülern mit geringer Kompetenz und Schülern mit sehr ausgeprägten Fertigkeiten mit zunehmendem Alter weiter wuchs: leistungsfähigere Schüler konnten demnach von ihrer bereits stärker ausgeprägten Lesefähigkeit profitieren und so den schon existierenden Vorsprung darüber hinaus noch ausbauen (vgl. Klicpera/Gasteiger-Klicpera 1993). Daher war die Bestrebung Leseleistungen bereits in frühem Alter zu diagnostizieren vorhanden, um dann unter Umständen Fördermaßnahmen zu veranlassen. Daher gibt es sehr viele Testverfahren zur Diagnostik der Leseleistung und Lesekompetenz im Bereich der Primarstufe. Für die Sekundarstufe bzw. für die Berufsschule jedoch existieren nur wenige deutschsprachige Tests (vgl. im Überblick Lenhard/Schneider 2009, S. 243f; Jordan 2011, S. 47).

Bei den angesprochenen Diagnostikinstrumenten für den Bereich der Primarstufe werden meist die Lesegeschwindigkeit und das Leseverständnis auf Wort-, Satz-, und Textebene versucht zu erfassen. Dazu werden beispielsweise Aufgaben zur Worterkennung, Lückentexten, oder Stolpertests, bei welchen in einem Satz ostentativ ein Wort zu viel enthalten ist, welches vom Leser identifiziert werden soll, gestellt. Dass sich diese Testverfahren nicht zum Einsatz in der Sekundarstufe oder der Berufsschule eignen, ist offenkundig, weswegen in den letzten Jahren – untermauert durch die negativen Resultate der PISA-Studien – vermehrt andere Lesetests für die Schüler der Sekundarstufe entwickelt wurden (vgl. Jordan 2011, S. 47).

Im Folgenden sollen in aller Kürze einige populäre Testverfahren kurz angerissen werden, welche für die Schülergruppe der beruflichen Schulen relevant sein könnten. Auf Grund unterschiedlicher sprachlicher Systematiken und gegebener Differenzen zwischen verschiedenen Sprachen werden im Folgenden nur die im deutschen Sprachraum gebräuchlichen Test bzw. die explizit in einer deutschsprachigen Variante vorliegenden berücksichtigt. Dies sind:

Der Frankfurter Leseverständnistest für 5. und 6. Klassen (FLVT 5-6)

Der seit 2008 eingesetzte *FLVT 5-6* wurde von Souvignier, Trenk-Hinterberger und Gold entwickelt. Dieser Test dient der Erfassung des Leseverständnisses von Schülern der Jahrgangsstufen fünf und sechs. Der Test wurde in zwei parallelen Testausführungen konstruiert, damit er als Gruppentest Einsatz finden kann und so Abschreibeffekte vermieden werden können. Dargeboten werden im *FLVT 5-6* zwei Texte mit jeweils ungefähr 570 Wörtern. Einer dieser Texte kann

als narrativer Text, der andere als Sachtext klassifiziert werden. Zu jedem Text werden 18 Items im Single-Choice-Format gestellt, d.h. jedes Item bietet vier Antwortmöglichkeiten, von denen exakt eine zutreffend ist. Der Test ist so konzipiert, dass er innerhalb einer Schulstunde als Einzel- und/oder Gruppentest durchgeführt werden kann. Dabei liegt die durchschnittliche Bearbeitungsdauer der Testfragen bei ca. 20 Minuten (vgl. Adam-Schwebe et al. 2009, S. 117). Bei der Entwicklung der in Summe 36 Testfragen wurden die von van Dijk und Kintsch beschriebenen Perspektiven des Leseverstehens antizipiert und umzusetzen versucht (vgl. Kintsch/van Dijk 1983). Daher wurden die Testfragen zu beiden Texten so entworfen, dass pro Hälfte textimmanente Verstehensleistungen zu erbringen sind und andererseits schlussfolgernde Verstehensleistungen abgefragt werden (vgl. Adam-Schwebe et al. 2009, S. 119). Dies sollte sicherstellen, dass Probanden einem bestimmten Fertigniveau des sinnentnehmenden Lesens zugeordnet werden können. Dazu werden zwei verschiedene Kompetenzstufen ermittelt. Probanden, welche die erste Kompetenzstufe erlangen, besitzen demzufolge die Kompetenz, eine oder mehrere ausdrücklich im Text genannte Informationen zu identifizieren und angemessene Schlussfolgerungen aus einer einzelnen Textpassage zu ziehen. Die zweite Stufe beinhaltet überdies die Fähigkeit, Informationen aus mehreren Textpassagen produktiv miteinander zu kombinieren oder, durch das in Beziehung setzen verschiedener Textpassagen Zusammenhänge herzustellen sowie die Fertigkeit, daraus Konklusionen ermitteln zu können (vgl. Adam-Schwebe et al. 2009, S. 122).

Dieser Test ist ein kommerzieller Test, bei dem auf Grund seiner Bekanntheit von einem relativ häufigen Einsatz ausgegangen werden kann. Bezüglich der Testgüte lässt sich festhalten, dass die interne Konsistenz (Cronbachs Alpha) der Testform A $\alpha = .88$ und der Testform B $\alpha = .86$ beträgt. Weiterhin beträgt die Paralleltestreliabilität der Testformen A und B $r_{tt} = .80$. Die Validität des FLVT 5-6 ist neben ihrer theoretischen Begründung auch empirisch belegt – bspw. korrelieren die ermittelten Werte der Schüler hoch mit den Leistungswerten in anderen Leseverständnistests ($r = .57$ bis $r = .74$). Des Weiteren liegen Normen für das zweite Halbjahr des Schuljahres 5 und 6 vor: Für die beiden Testformen liegen z- und T-Äquivalenzwerte für die fünften und sechsten Jahrgangsstufen der verschiedenen Schularten vor (Form A: $N = 1.239$; Form B: $N = 1.237$) (vgl. Souvignier et al. 2009).

Das Salzburger Lese-Screening für die Klassenstufen 5-8 (SLS 5-8)

Von Auer, Gruber, Mayringer und Wimmer wurde im Jahr 2005 das SLS 5-8 entwickelt. Dieser Test orientiert sich an den bereits zuvor entworfenen Salzburger Lese-Screenings der Klassenstufen 1-4. Das SLS soll eine strukturierte Diagnose von Leseschwächen ermöglichen und Differenzen in der basalen Lesefähigkeit ökonomisch wie auch akkurat erfassen (vgl. Schneider 2009, S. 56). Im SLS steht eher der technische Aspekt des Lesens im Fokus. Das meint, dass bezugnehmend auf die kognitionspsychologische Leseforschung, welche aktuell das Wortlesen als wesentliche Komponente des Leseprozesses identifiziert hat, wird im SLS das schnelle Lesen einfacher Sätze altersadäquat geprüft. Das Vorgehen ist dergestalt: innerhalb einer Bearbeitungszeit von drei Minuten müssen Probanden eine Liste mit 70 inhaltlich einfachen Sätzen möglichst schnell lesen und den Wahrheitsgehalt der Sätze überprüfen. Für jeden innerhalb dieser

Zeit korrekt beurteilten Satz wird ein Punkt vergeben. Das Testergebnis setzt sich sodann aus der Summe aller richtig bewerteten Sätze zusammen (vgl. Auer et al. 2005a). Somit lässt sich feststellen, dass es primär darum zu gehen scheint, die Lesegeschwindigkeit zu erfassen. Bezüglich der Testgüte ist festzuhalten, dass auch dies ein kommerzieller und empirisch geprüfter Test ist. Dabei beträgt die angegebene Reliabilität $r = .89$ und kann ebenfalls als hoch eingeordnet werden. Die Validität liegt bei $r = .78$. Ferner traten bei der Erprobung des Verfahrens keine Boden- oder Deckeneffekte auf, d.h. der Test differenziert jeweils über den gesamten Leistungsbereich ohne Einschränkungen. Für jede der vier Schulstufen liegt eine separate Normierung vor, die auf Stichproben zwischen 714 und 850 Schülerinnen und Schülern basiert.

Der Lesegeschwindigkeits- und -verständnistest für die Klassen 6-12 (LGVT 6-12)

Der von Schneider et. al im Jahr 2007 für Probanden im Altersbereich von 11-18 Jahren entwickelte *LGVT 6-12* stellt momentan eines der wenigen standardisierte Gruppentestverfahren zur Ermittlung der Lesekompetenz auch von Schülern der Jahrgangsstufen 9 bis 12 dar. Dieser Test ist als kurzer Lückentext mit Auswahlmöglichkeiten konstruiert worden. Es werden zur Komplettierung eines Satzes mehrere Optionen geboten. Durch das Selektieren der richtigen Vervollständigungsalternative wird theoretisch verhindert, dass der Text nur oberflächlich gelesen wird. Der Text ist dabei bezüglich seiner Länge so angelegt, dass innerhalb der Bearbeitungszeit kein Proband in der Lage sein soll, ihn vollständig zu lesen. Die Bearbeitungszeit beträgt insgesamt vier Minuten. Wenn diese Zeit abgelaufen ist, müssen die Schüler einen Strich hinter das letzte gelesene Wort machen, so dass somit neben dem Wert für jedes richtige Wort ein Wert für die Lesegeschwindigkeit angegeben werden kann. Dieses Verfahren bietet folglich die Möglichkeit, neben dem Leseverständnis auch die Lesegeschwindigkeit zu messen (vgl. Schneider 2009, S. 58). Bezüglich der Testgüte ist festzuhalten, dass die Retestreliabilität (nach 6 Wochen) bezüglich des Leseverständnisses bei $r = .87$ ($N = 103$) und bezüglich der Lesegeschwindigkeit bei $r = .84$ ($N = 103$) liegt. Der *LGVT 6-12* weist beim Leseverständnis eine Korrelation von $r = .59$ ($N = 711$) zum PISA2000-Leseverständnistest und zum Lesestrategiewissen WLST eine Korrelation von $r = .46$ ($N = 809$) auf, jedoch nur eine geringe Übereinstimmung zum Gates/MacGinitie (vgl. Norwig et al. 2013). Die Normierungsstichprobe setzte sich aus insgesamt 2.390 Schülern aus elf deutschen Bundesländern zusammen. Die diesbezüglich ermittelten Normen für die Anzahl korrekter Unterstreichungen sowie die Menge der gelesenen Wörter liegen für die Klassenstufen 6-9 der allgemeinbildenden Schularten vor. Darüber hinaus sind noch Normen für Realschule 10. Klasse sowie Gymnasium Klassenstufe 10 und 11 verfügbar (vgl. Auer et al. 2005b).

Der Würzburger Lesestrategie-Wissenstest für die Klassen 7-12 (WLST 7-12)

Der *WLST 7-12* für die Klassenstufen 7-12 wurde von Schlagmüller und Schneider entwickelt und im Jahr 2007 veröffentlicht. Dieser ist ebenfalls sowohl als Einzel- oder auch Gruppentest allerdings in Papier- oder PC-Version durchführbar. Seine Bearbeitungszeit beträgt annähernd 30 Minuten. Die Lesekompetenz wird dabei nicht auf der Textebene erfasst, sondern viel mehr

auf der Prädiktorebene über das Lesestrategiewissens, welches sicher auch einen Teil einer ausgeprägten Kompetenz ausmacht (siehe PISA-Befunde). Es soll demzufolge durch den WLST 7-12 überprüft werden, ob Probanden zumindest nominell über effektive Strategien zum Verarbeiten und Verstehen von gelesenen Texten verfügen, d.h. es wird erfasst, ob sie derartige Strategien kennen und beurteilen können – über deren tatsächliche Anwendung können allerdings keine Aussagen getroffen werden. Es werden den Probanden sechs verschiedene Lese- und Lernszenarien präsentiert. Zu diesen Szenarien werden jeweils fünf verschiedene Verfahren zum Erreichen des jeweiligen Ziels genannt. Die Probanden sollen sodann die präsentierten Vorgehensweisen in einer Abfolge anordnen (das ihrer Ansicht nach optimale Vorgehen an erster Stelle, das unvorteilhafteste an fünfter Stelle). Die resultierende Klassifizierung bzw. das Ranking bezüglich der Vorgehensweisen wird der Anordnung von Experten gegenübergestellt. Eine hohe Übereinstimmung resultiert dabei in einem hohen Testwert. Die verwendete Rangfolge der Experten wurde im Vorfeld durch Interviews von Lehrkräften und pädagogischen Psychologen erstellt (vgl. Schlagmüller/Schneider 2007b). Hierbei beträgt die Split-Half-Reliabilität $r = .88$, die interne Konsistenz: $r = .88$ und die Retestreliabilität (nach 6 Wochen): $r = .81$. Bezüglich der Gültigkeit kann im Bezug zum Leseverständnis nach PISA 2003 eine Korrelation von $r = .40$ ($N = 3.386$) und zum Schnellleseverständnistest (SLVT) eine von $r = .46$ ($N = 809$) festgehalten werden. Es liegen Normen (Normierungsstichprobe $N = 4.490$ Schülern aus allen deutschen Bundesländern) für die erreichte Punktzahl aus den Paarvergleichen vor (vgl. Schlagmüller/Schneider 2007a).

Zusammenfassend lassen sich Erkenntnisse (vgl. Jordan 2011) bestätigen, dass erst in den letzten Jahren verstärkt Tests für die Sekundarstufe zur Verfügung gestellt wurden. Die bisher konstruierten und oben vorgestellten Tests decken jedoch nicht umfassend das Spektrum der Lesekompetenz ab. Bei diesen bisher beschriebenen Testverfahren werden ausschließlich kontinuierliche Texte in Form von narrativen Texten oder Sachtexten eingesetzt. Keiner der oben aufgeführten Tests zur Ermittlung der Leseleistung und Lesekompetenz verwendet systematisch depiktionale Repräsentationen.

In den Kapiteln 2.1.2 und 3.1.3 wurde jedoch ausführlich dargelegt, dass Leser im beruflichen und berufsbildenden Bereich insbesondere mit verschiedenen Formen deskriptionalen und depiktionalen Repräsentationsformats wie Karten, Diagrammen, Anleitungen, Bilder, Pläne und Tabellen konfrontiert werden. Da diese Anforderungen in keinem der genannten Testverfahren abgebildet wurden, lohnt es sich Lesetests speziell für Berufsschüler und Erwachsene zu betrachten. Dazu gehören einerseits der LTB3, die ULME Untersuchungen und auch PIAAC.

Der Lesetest für Berufsschüler/innen LTB-3

Der *LTB3* (vgl. Drommler et al. 2006) greift das Kompetenzstufenmodell aus PISA auf, modifiziert es jedoch stark – aus dem ursprünglich dreiskaligen Modell in PISA wird ein Konzept mit nur einer Skala – welche in drei Kompetenzstufen eingeteilt ist. Es wird wie folgt modelliert:

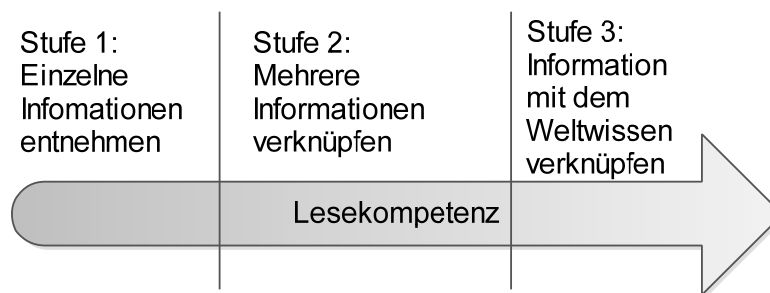


Abbildung 8 – Kompetenzmodell des LTB3

Die erste Kompetenzstufe umfasst die Fertigkeit zur expliziten Informationserfassung eines Textes und die korrekte und umfassende Wiedergabe dieser. Auf der zweiten Kompetenzstufe werden über diese basale Fähigkeit hinaus die Analyse relevanter Informationen und deren Zusammenhänge im Hinblick auf einen spezifischen Auftrag vom Probanden gefordert. Dazu gehören die eigenständige Selektion und Analyse von Konvergenzen, Divergenzen und Entwicklungen, sowie die Strukturierung der zentralen Informationen und der Konstruktion von Transfermöglichkeiten. Die dritte Kompetenzstufe zielt auf die Bewertung der relevanten Kerninhalte der Texte, Tabellen und der Grafik seitens der Probanden und das in Beziehung setzen zum eigenen Weltwissen. Diese Fähigkeit erst ermöglicht laut Drommler et al. ein kritisches Hinterfragen eines Textes sowie der individuellen Position und einen Transfer zur Lösung einer konkreten Aufgabe (vgl. Drommler et al. 2006, S. 14f).

Die genannten drei Kompetenzstufen bauen hierarchisch aufeinander auf, d.h. ohne singuläre Informationen aus einem Text entnehmen zu können (Stufe 1), ist es nicht möglich, verschiedene Informationen zu integrieren (Stufe 2). Gleichfalls ist es dann nicht möglich, diese Informationen mit dem Weltwissen zu verknüpfen und dadurch zu einer durchdachten Auffassung oder einer verlangten Lösung zu gelangen (Stufe 3). Somit stellt eine Kompetenzstufe eine notwendige Voraussetzung zur Erlangung der nächsten Hierarchiestufe dar (vgl. Drommler et al. 2006, S. 15).

Im LTB3 werden ausschließlich non-fiktionale Texte dargeboten und vergleichbar zur PISA-Test Konstruktion werden lineare und non-lineare Repräsentationsformate vorgelegt. Um darüber hinaus die Motivation der Berufsschüler sicherzustellen, werden vergleichsweise reale Texte, Tabellen und Grafiken aus der angenommenen Lebenswelt der Schüler gewählt. Diese wurden während der Konstruktion nur wenig – vornehmlich im Layout (Tabellen und Grafik) – modifiziert. Dabei werden sowohl zu Texten, als auch zu Diagrammen und Tabellen Fragen in jeder Kompetenzstufe gestellt. Nach einer Item analytischen Beurteilung kamen Drommler et al. allerdings zu der Erkenntnis, dass eine separate Auswertung jedes Repräsentationsformats keine interpretierbaren Ergebnisse liefert, da die Anzahl der Fragen zu gering sei (vgl. Drommler et al. 2006, S. 16; 29). Weiterhin wissenswert ist, dass der Test ein sogenannter Niveau-Test ist, das bedeutet, dass allen Schülern genügend Zeit zur Verfügung gestellt wird, um alle Items zu bearbeiten. Das Testformat ist single-choice Paper-Pencil, bei dem es zu allen Fragen nur eine richtige Antwort gibt (vgl. Drommler et al. 2006, S. 23).

Als Messverfahren für den Einsatz in der beruflichen Bildung versucht der Test die im Folgenden genannten spezifischen Anforderungen und Rahmenbedingungen der Berufsschule zu berücksichtigen (vgl. Drommler et al. 2006, S. 13; 16):

- Konzentration auf wesentliche in der Berufswelt und in der Ausbildung relevante Textsorten
- Offenlegung der darauf bezogenen individuellen Stärken und Förderbedarfe hinsichtlich der Lesekompetenz der einzelnen Schülerin und des einzelnen Schülers
- Praktikabilität im Hinblick auf den Umfang und die methodische Durchführung des Diagnoseverfahrens im Unterrichtsalltag
- Bereitstellung von möglichst konkreten Diagnoseergebnissen zur Lesekompetenz, die über den Unterricht im Fach Deutsch/Kommunikation hinausgehend Möglichkeiten für eine fächerübergreifende Leseförderung bieten.

Die von Jordan (2011) durchgeführten Analysen geben allerdings Anlass, die Messqualität des Tests respektive einiger Items zu bezweifeln (vgl. Jordan 2011, S. 175f).

Untersuchung der Leistungen, Motivation und Einstellungen zu Beginn der beruflichen Ausbildung (ULME)

ULME kann als eine Fortführung der etablierten Längsschnittstudie „Aspekte der Lernausgangslage und der Lernentwicklung“ (LAU) für den Bereich der berufsbildenden Schulen gesehen werden. Die LAU Untersuchungen sind im Schuljahr 1996/97 mit den derzeitigen fünften Klassen verschiedener Hamburger Schulen erstmalig durchgeführt worden. Die Behörde für Bildung und Sport (BBS) in Hamburg gab Anfang des Jahres 2002 eine wissenschaftliche Untersuchung bezüglich der Leistungen, Motivation und Einstellungen von Schülern des beruflichen Bildungssystems in Hamburg im Schuljahre 2002/03 in Auftrag. Die ULME-Studie entstand somit zusammenfallend mit der letzten Phase der Längsschnittuntersuchung „Aspekte der Lernausgangslage und Lernentwicklung“ (LAU 11), welche vollständig Hamburger Fünftklässler des Schuljahres 1996/97 aus staatlichen Schulen einbezieht und deren erreichte Lernstände, Lernentwicklungen und schulbezogenen Einstellungen vom Ende der Grundschulzeit bis in die Klassenstufe 11 erfasst. Ein großer Teil der in die ULME untersuchten Probanden gehörte auch zur LAU-Kohorte (vgl. Lehmann et al. 2005, S. 5).

Zentral für die gesamten ULME-Studien sind Fragen und Untersuchungen bezüglich der Aspekte der Lernausgangslagen in den Eingangsklassen der Beruflichen Schulen. Dabei werden individuelle und soziale Bildungsvoraussetzungen sowie auf Berufsbildungsentscheidungen und -barrieren und deren Entwicklungen bis zu ULME III beobachtet und analysiert. Ziel der Studien war es durch die Ergebnisse der Untersuchungen dem Amt für Bildung, Abteilung Berufliche Bildung und Weiterbildung eine optimierte Basis für die Beratung und Unterstützung der beruflichen Schulen zur Verfügung zu stellen (vgl. Lehmann et al. 2005, S. 5).

Dazu wurden für die Untersuchung zunächst standardisierte Fachleistungstests, ein spezielles Verfahren zur Ermittlung von Textverarbeitungswissen, ein Fragebogen zu Lernstrategien (Vorläufer des WLST), Aufgaben zur Textproduktion sowohl für das Fach Deutsch als auch für das Fach Englisch sowie ein Schülerfragebogen adaptiert/konzipiert und im Herbst 2002 eingesetzt. Auf Grund der ausgeprägten Heterogenität der Zusammensetzung der Schüler der Beruflichen Schulen bezüglich des Schulabschlusses – teilweise ohne Hauptschulabschluss bis hin zu Absolventen mit Hochschulzugangsberechtigung – wurden die Tests zur Lesekompetenz in drei unterschiedlichen Versionen entsprechend den angesprochenen Schulabschlussniveaus (Hauptschule (H), Realschule (R) und Gymnasium (Gy)) konstruiert (vgl. Lehmann et al. 2005, S. 7).

In der ersten Phase der Untersuchung (ULME I) standen vor allem deskriptionale Texte im Fokus. Dabei enthält das nach Anforderungsniveau differenzierte Design insgesamt 32 mehrheitlich im Multiple-Choice-Format gestellte Aufgaben. Den Items zum Leseverstehen liegt eine Testkonzeption zugrunde, welche bereits in der LAU-5- und in der LAU-7-Untersuchung angewandt wurde. In jeder Testversion werden Items verschiedener Schwierigkeiten eingesetzt, so dass der ganze Anforderungsbereich angefangen von simpler Informationsidentifikation bis hin zur selbständig schlussfolgernden bzw. interpretatorischen Arbeit mit Texten erfasst werden kann. Insgesamt gab es 32 Items pro Testversion zu vier Texten, d.h. zu jedem Text lagen 8 Aufgaben vor. Von diesen vier Texten gehörten jeweils zwei in den längsschnittlichen Ankerbereich⁷⁶, einer in den querschnittlichen Kernbereich und einer im nach Anforderungsniveau (Schulabschluss) spezifizierten Bereich (vgl. Lehmann et al. 2005, S. 10).

Die beiden Ankertexte mit ihren zusammen 16 Ankeritems, dienen als Berechnungsgrundlage für Längsschnittanalysen. Sie stammen aus dem im Jahr 2000 eingesetzten Hamburger Schulleistungstest für achte und neunte Klassen. Die kern- und testversionsspezifischen Texte wurden neu ausgewählt und nach einer Pilotierung mit 1.510 Schülern im Herbst 2001 für die Hauptuntersuchung eingesetzt. Die Tests wurden jeweils in Parallelformen entwickelt, so dass eine A-Form und eine entsprechende B-Form ausgegeben werden konnte, wobei sich diese Alternativen lediglich in der Reihenfolge der Texte unterscheiden. Die Bearbeitungszeit betrug inklusive der notwendigen organisatorischen Einführung in den Test ca. 60 Minuten (vgl. Lehmann et al. 2005, S. 10).

Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der teilqualifizierenden Berufsfachschulen (ULME II) und Untersuchung von Leistungen, Motivation und Einstellung der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen (ULME III)

Die oben genannte Studien werden fortgeführt in *ULME II* (Lehmann et al. 2006) und *ULME III* (Lehmann/Seeber 2007). Diese haben als zentrales Anliegen und besondere Herausforderung empirisch fundierte Aussagen zu den am Ende einer Berufsausbildung im Rahmen des

⁷⁶ Stark vereinfacht: Set von gemeinsamen Items („Anker-Items“), die verschiedenen Stichproben/Probandengruppen/Subgruppen dargeboten werden, so können Test- oder Fragebögen auf spezifische Gruppen abgestimmt werden, bleiben aber durch/mit eine/r gemeinsame/n Skala verknüpft.

dualen Systems erreichten allgemeinen und beruflichen Kompetenzen zu treffen. ULME III erfasst daher im Rahmen des Tests „Texte und Tabellen“ sogenannte allgemeine berufliche Grundqualifikationen, welche als grundlegende Bedingungen für die Gestaltung der individuellen Handlungsoptionen im Alltag sowie für die aktive Partizipation am beruflichen und gesellschaftlichen Leben gelten können (vgl. Lehmann et al. 2007, S. 16).

Für den Test zu den beruflichen Grundqualifikationen, was die basalen Lese- und Rechenfertigkeiten meint, konnte sich auf Items der Internationalen Studie zu den Grundqualifikationen der erwerbsfähigen Bevölkerung (International Adult Literacy Survey: IALS) gestützt werden. In der Hauptsache wurden bezüglich Lesekompetenz folgenden Fähigkeiten der Verarbeitung mit unterschiedlich präsentierten und verschieden strukturierten Informationen analysiert (vgl. Lehmann/Hunger 2007, S. 24):

1. Kompetenzen im Umgang mit kontinuierlichen Texten (bezeichnet als „prose literacy“), d.h. Informationsidentifikation und -entnahme aus alltäglichen oder allgemein beruflichen Texten, wie sie z.B. in Zeitungen/Zeitschriften oder im Informationsmaterial von Betrieben enthalten sind
2. Kompetenzen im Umgang mit diskontinuierlichen Texten (bezeichnet als „document literacy“), d.h. Informationsidentifikation und -entnahme aus Tabellen, grafischen Präsentationen und schematischen Darstellungen, wie sie sowohl im Kontext von beruflichen Tätigkeiten als auch im Freizeitbereich, im Agieren als Konsument, oder auch in Presseerzeugnissen auftreten

In der ursprünglichen Studie (IALS) wiesen die eingesetzten Tests ein offenes Antwortformat auf und zeichneten sich durch den Einsatz authentischer – oder zumindest ebendiesen Anschein erweckender – Stimuli (wie bspw. Zeitungsausschnitte, Formulare, etc.) aus. Insgesamt konnten dort 114 Items in Form eines Balanced-Incomplete-Block-Designs eingesetzt werden, wobei die Tests mit externen Testleitern in Form von Interviews durchgeführt wurden. Dieses verhältnismäßig aufwendige Verfahren konnte im Rahmen von ULME III aus erhebungs- und auswertungsökonomischen Gründen nicht entsprechend eingesetzt werden. Aus diesem Grund mussten zum Teil die offenen Antwortformate in ein Multiple-Choice-Format transformiert werden. Dies betraf vornehmlich die Items zur Informationsentnahme aus kontinuierlichen Texten und in einigen Fällen auch Aufgaben zu den diskontinuierlichen Texten. Bei den Items bezüglich der fundamentalen Rechenfertigkeiten konnte grundsätzlich das offene Antwortformat beibehalten werden, so dass das Datenmaterial später einer etwaigen Fehleranalyse zugänglich gemacht werden kann. Von den insgesamt 114 Aufgaben wurden 41 gewählt. Durch die angesprochenen Formatänderungen wurden 62 Variablen – davon 22 wie beschrieben in das Multiple-Choice-Format transformiert – erhoben. Dieser Test hatte eine durchschnittliche Bearbeitungszeit von 75 Minuten (vgl. Lehmann/Hunger 2007, S. 24f).

Programme for the International Assessment of Adult Competencies (PIAAC)

Mit *PIAAC* analysiert die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) zentrale Grundkompetenzen in der erwachsenen Bevölkerung. Untersucht werden unter anderem die Lesekompetenz, die alltagsmathematische Kompetenz und technologiebasiertes Problemlösen – Fähigkeiten von denen angenommen wird, dass sie für die erfolgreiche Teilhabe an der heutigen Gesellschaft von zentraler Bedeutung sind. Diese Kompetenzen werden in der Forschung als ein wichtiger Grundstein für die Entwicklung zahlreicher weiterer spezifischer Kompetenzen und Fertigkeiten gesehen. Durch *PIAAC* soll die Möglichkeit geboten werden diese Schlüsselkompetenzen im Erwachsenenalter international vergleichen zu können. Das Augenmerk liegt in diesem Fall auf der Bevölkerung im erwerbsfähigen Alter, d.h. von 16 bis 65 Jahre. Somit versucht *PIAAC* Informationen darüber zu liefern, inwieweit die erwerbsfähige Bevölkerung in den jeweiligen Ländern in den untersuchten Grundkompetenzen divergiert. Zudem sollen Faktoren untersucht werden, welche mit dem Erwerb und dem Erhalt dieser Kompetenzen in Zusammenhang stehen (vgl. Rammstedt 2013, S. 11).

In *PIAAC* wird Lesekompetenz definiert als das Verstehen, Interpretieren sowie Nutzen von geschriebenem Material und dabei verstanden als die Voraussetzung, um das eigene Wissen und Potenzial weiterzuentwickeln und am gesellschaftlichen Leben teilzunehmen. Um den Realitätsbezug abzubilden sind in *PIAAC* Aufgaben wie das Lesen und Verstehen eines Medikamentenbeipackzettels oder eines kurzen Zeitungsartikels enthalten. Daneben gibt es Aufgaben, die auch elektronische Medien tangieren, wie beispielsweise das Lesen einer Stellenanzeige in einem Onlineportal (vgl. Rammstedt 2013, S. 12). Zur Messung der Basiskompetenzen wurden gemäß der jeweiligen theoretischen Konzeption der Kompetenzdomäne entsprechende Aufgaben entwickelt, welche basierend auf Modellen der Item-Response-Theorie (siehe Kapitel 0) skaliert wurden. Dabei wurde jede Kompetenzdomäne auf einer eigenen Skala modelliert. Zur besseren Interpretation und zur Beurteilung der Werte wurden diese Skalen äquivalent zu vergleichbaren Studien (bspw. PISA) jeweilig in Kompetenzstufen mit Intervallen von je 50 Kompetenzpunkten gegliedert. Daraus ergeben sich fünf Kompetenzstufen für die Lesekompetenz, wobei darüber hinaus auch der Bereich unterhalb der niedrigsten Stufe klassifiziert wird („Unter Stufe I“) (vgl. Rammstedt 2013, S. 12).

PIAAC wurde bereits 2008 von den OECD-Mitgliedsstaaten initiiert und soll weiterhin ähnlich wie PISA in periodisch wiederkehrenden Zyklen repetiert werden. Die stattgefundenen Erhebungen in 2012 stellen daher in der Konzeption lediglich den Startpunkt dar, wiederholte Erhebungswellen sind in einem 10-Jahres-Turnus geplant sind (vgl. Rammstedt/Zabal 2013, S. 21). Der Test ist so konzipiert, dass die Probanden ihn selbstständig unter Anwesenheit des Interviewers am Computer oder in Papierform bearbeiten können (vgl. Rammstedt 2013, S. 13).

Gates-MacGinitie-Test bzw. dessen kultursensible Übersetzung

Weiterhin existiert eine Übersetzung des standardisierten US-amerikanischen Gates-MacGinitie-Tests (der Originalstudie von Palincsar/Brown (1984)), ins Deutsche (vgl. Petsch et al.

2008; Gschwendtner 2012). Dieser Test zur Erfassung des Leseverständnisses liegt für unterschiedlichen Klassenstufen (bspw. 7/9 oder 10/12) vor. Getestet wird das Verständnis kontinuierlicher Texte. Bezüglich der Adressaten und des hier vorliegenden Kontextes (BVJ) wurde von Gschwendtner und hinzugezogenen Fachdidaktikern der Gates-MacGinitie 7/9 als passend und adressatentauglich (bezüglich Kompetenz- und Altersangemessenheit) befunden (vgl. Gschwendtner 2012, S. 139).

Im Original liegt der Test in parallelisierten Testformen vor und gehört zur Kategorie der sogenannten high-stakes Tests, d.h. dass die Testgütekriterien strengen Maßstäben gerecht werden. Die Reliabilität des Originaltests⁷⁷ liegt bei .93. Beide vorliegenden Testformen umfassen jeweils 11 Kurztexte (Länge zw. 71 und 171 Wörtern) und sowohl literarische als auch Sachtexte mit 48 Items im Multiple-Choice-Format mit jeweils vier Antwortalternativen von denen nur jeweils eine richtig ist (vgl. Gschwendtner 2012, S. 139).

Bei der Übersetzung dieses Tests – mit der Absicht und dem Ziel die hohen Testgütekriterien ebenfalls zu transferieren – ist besondere Sensibilität bezüglich kultureller und sprachlicher Besonderheiten notwendig. Insbesondere bei Redewendungen und speziellen syntaktischen Merkmalen muss besondere Sorgfalt bewiesen werden, da bei einer allzu wörtlichen Übersetzung die Texte geradezu entstellende und komikartige Züge annehmen können. Dennoch können auch qualitativ hochwertige und mit großer Expertise angefertigte Übersetzungen (bspw. die PISA Test) nicht als frei von Verzerrungen bezeichnet werden. Die Übersetzung des Tests hatte einen holistischen Transfer zum Ziel, welche die inhaltliche Validität sicherstellt. Dies wurde durch die gemeinsame Übersetzung des Tests durch zwei Personen und eine weitere parallele Übersetzung einer dritten Person, welche anschließend über die Ergebnisse diskutierten und z.T. neue Möglichkeiten entwickelten und abschließend zu einer gemeinsamen (optimierten) Fassung kamen, gewährleistet. Weiterhin erfuhr diese Fassung durch einen Experten sprachlichen Feinschliff und wurde alsdann von einem Lektor wieder ins Englische transferiert. Die sich ergebenden und erwarteten Abweichungen wurden schließlich gemeinsam besprochen und für weitere Optimierungen genutzt. Die so entstandenen Tests wurden in einem letzten Schritt parallelisiert (vgl. Gschwendtner 2012, S. 139f).

Die Itemschwierigkeiten korrelieren mit denen des Originals mit $r=.81$, was insgesamt diesbezüglich für einen hohen Testgüte spricht. Die Bearbeitungszeit beträgt 39 Minuten. Die Items wurden dichotom gescored und anschließend mittels eines einparametrischen, eindimensionalen, dichotomen Raschmodells geschätzt (vgl. Gschwendtner 2012, S. 140). Trotz der anzunehmenden Güte des Tests und der erfolgten Übersetzung, kann dieser auf Grund datenschutzrechtlicher Bedenken des Verlags Riverside Publishing nicht veröffentlicht werden (vgl. Gschwendtner 2012, S. 140), so dass von einer Nutzung an dieser Stelle abgesehen werden muss.

Bei der Reflexion der Passung auf den beruflichen Kontext zeigt sich, dass vor allem die ULME Konzeption und die PIAAC Items auf Grund ihrer Berücksichtigung verschiedener Repräsenta-

⁷⁷ Basis: Stichprobe der amerikanischen 9. Klasse

tionsformate – depiktionaler wie deskriptionaler Natur – den Anforderungsbereich des täglichen beruflichen Umgangs mit Texten widerspiegeln. In keiner dieser Testkonzeptionen werden allerdings die Lesesituation bzw. der Anforderungskontext und deren möglicher Einfluss auf die Kompetenzausprägung beachtet. Auch werden die verschiedenen Repräsentationsformate zwar als Präsentationsformen der Stimuli benutzt, jedoch bleiben die diesbezüglichen Analysen größtenteils oberflächlich. Eine Ausnahme stellt hierbei ULME dar. Hier wurden entsprechende Analysen durchgeführt, die einen Unterschied zwischen kontinuierlichen und diskontinuierlichen Texten ergaben.

Aus den genannten Gründen und der Tatsache, dass die meisten Tests nicht für den Einsatz verfügbar sind bzw. sich nicht spezifisch auf die hier interessierenden Fragestellungen⁷⁸ und insbesondere den beruflichen Kontext zugeschnitten sind, unterbleibt eine 1:1 Übernahme aller Items aus einem Test.

3.1.7. Rekapitulation des Untersuchten II

Zusammengefasst lassen sich folgende zentrale Aussagen festhalten:

1. Der Begriff Lesen ist so vielschichtig wie im Detail umstritten.
2. Bezüglich des vermutlich nicht unerheblichen Einflusses der Lesesituation bzw. des Anforderungskontextes gibt es bisher nur wenige empirische Belege, jedoch viele theoretisch fundierte Hinweise.
3. Die Kommunikationssituation beim Lesen insgesamt ist kognitiv höchst anspruchsvoll und keineswegs trivial.
4. Prozesse beim arbeitsplatzbezogenen Lesen (Lesen im Handlungskontext) differieren vermutlich von schulbezogenen Leseprozessen (Lesen im Lernkontext).
5. Das typische berufliche Lesematerial ist vielschichtig und besitzt verschiedene Repräsentationsformate sowie fragmentarischen Charakter
6. Es gibt einige Leseprozesstheorien, die verschiedene Ebenen differenzieren. Sie fokussieren jedoch zumeist deskriptionale Texte. Nicht unumstritten, trotzdem jedoch am plausibelsten, scheint zum gegenwärtigen Zeitpunkt, dass Textverstehen den Aufbau eines Mentalen Modells voraussetzt, zumindest sofern die Nutzung der Informationen in das Konstrukt des Textverstehens einbezogen wird.

⁷⁸ Siehe Kapitel 0 – S. 211.

7. Bezüglich der Verarbeitung von depiktionalen Strukturen gibt es weit weniger Betrachtungen. Auch diese kommen ebenfalls zu einer Theorie, die vom Aufbau eines Mentalen Modells ausgeht. Dies ist auf Grund der analogen Prozesse naheliegend.
8. Es gibt zwei zentrale Theorien des gemeinsamen Bild-und-Text-Verständnisses – die Dual-Coding-Theory (vgl. u.a. Sadoski/Paivio 2004) und das integrative Text-Bild-Verstehen (vgl. u.a. Schnotz 1997). Beide Ansätze postulieren die Konstruktion eines Mentalen Modells und kognitive Verarbeitungsprozesse bei einem Moduswechsel.
9. In der Textverständlichkeitsforschung finden sich fast ausschließlich Betrachtungen deskriptionaler Repräsentationen. Was die Verarbeitung logischer oder realer Bilder erschwert oder erleichtert ist theoretisch wenig erörtert und empirisch noch weniger belegt. Insgesamt kann als ein Ergebnis der Textverständlichkeitsforschung bilanziert werden, dass sich die Textverständlichkeit kaum unabhängig von der Anforderungssituation und dem Leser bestimmen lässt (vgl. auch dazu das Tetra-Eder-Modell von Jenkins, S. 48).
10. Die Lesekompetenzdiagnostik hat sich in den letzten Jahren sehr stark entwickelt. Fokussierte sie zu Beginn noch Schwächen auf basalen Ebenen, gibt es heute ein überaus komplexes Verständnis des Leseverstehens und verschiedenen Messinstrumente, um es zu erfassen.
11. Optimal passend für den beruflichen Bereich ist noch kein Instrument, jedoch sind insbesondere in den letzten fünf Jahren starke Tendenzen zu beobachten diese Lücke zu schließen. Eine Berücksichtigung der Lesesituation und des Anforderungskontextes ist jedoch bisher nicht zu verzeichnen.

Bezüglich des 11. Punktes schließt sich diese Arbeit an. Im Folgenden wird der Rahmen für die Entwicklung eines Testinstruments erläutert und die Konstruktion ebendieses Tests vorgestellt. Daran schließen sich erste Auswertungen der entwickelten Items und eine kritische Reflexion an.

4. Ansätze zur Entwicklung eines Instruments zur Untersuchung des Einflusses der Repräsentationsformate und kognitiven Anforderungsbereiche

4.1. Entwicklung des Testinstruments

4.1.1. Überlegungen zum Testformat

Bevor ein Instrument zur Überprüfung beruflicher Lesekompetenz entwickelt werden kann, gilt es, sich Gedanken um das Testformat zu machen. Teigeler (1968) hat verschiedene grundsätzliche Testformate gegeneinander abgewogen, ohne sich dabei auf spezielle Lesetests zu beziehen:

Die einfachste Methode den Grad des Behaltens oder Verstehens⁷⁹ von Texten zu erforschen ist, laut Teigeler, Probanden einen Text zu geben, sie nach der Lektüre zu bitten das Behaltene aufzuschreiben und diese Niederschrift anschließend in Klassen einzuteilen, so dass sie quantifizierbar sind. Durch zählen der Klassenhäufigkeiten kann festgestellt werden, wie viel des Behaltenen verstanden wurde. Problematisch an dieser Vorgehensweise ist die objektive Bewertung und Klassifizierung des Materials. Ferner ist diese Methode sehr zeitaufwändig, sodass sie kaum noch praktiziert wird (vgl. Teigeler 1968, S. 54). Auch zu bezweifeln ist, ob das Wiedergeben von Textfragmenten/-bausteinen tatsächlich auf eine höhere kognitive Verarbeitung und damit eine Verstehensleistung schließen lässt – dies darf mit großer Wahrscheinlichkeit bezweifelt werden.

Ein anderes Verfahren ist der sogenannte Ratetest. Hierbei muss der Proband jeden Buchstaben eines Textes so lange erraten, bis der richtige gefunden wurde. Es wird mit nur einem einzelnen Buchstaben begonnen und sukzessive bei richtigem Raten ergänzt, so dass insbesondere zu Anfang noch sehr viele Fehler gemacht werden. So soll der subjektive Informationsgehalt eines Textes ermittelt werden können. Kritisch an diesem Verfahren ist einerseits die Dauer (ca. 1 - 2h für 300 Wörter), so dass entweder nur sehr kurze Texte genommen werden können oder die Probanden über Gebühr belastet werden (vgl. Teigeler 1968, S. 55f).

Deutlicher wird dies an einem Beispiel: Gegeben ist der Buchstabe „I“, der Proband wird nun aufgefordert den nächsten Buchstaben zu erraten. Denkbare Versuche sind bspw. c, n, oder m. Da „m“ der richtige Buchstabe ist, wird dieser ergänzt, so dass dort steht „Im“, jetzt muss der Proband erraten, ob als nächstes ein weitere Buchstabe (bspw. für das Wort „immer“ oder eine Lücke (für das Wort „Im“) folgt. In diesem Beispiel folgt eine Lücke, wie viele Versuche der Proband benötigt, sollen Hinweise darauf sein, wie gut der Text verstanden wird. Auf „Im“ kann fast alles folgen, so dass der Proband möglicherweise alphabetisch beginnt und hier das „A“⁸⁰ richtig errät. „Im A“ lässt immer noch sehr viel Raum für Interpretation und zufälliges – statt strategisches – Raten. Bis der Satz „Im Anfang war das Wort und das Wort war bei Gott, und

⁷⁹ Eine Gleichsetzung die in der Literatur häufiger vorgenommen wird. Ihre Zulässigkeit darf jedoch bezweifelt werden, da Erinnerungsleistungen grundsätzlich von Verstehensleistungen abzugrenzen sind, auch wenn in der Forschung ein gewisser Zusammenhang (These: was verstanden wurde, kann besser behalten werden) angenommen wird (Naceur 2001, S. 113f).

⁸⁰ Angemerkt werden muss: zufällig!

das Wort war Gott.“ erraten wurde dürfte einige Zeit vergehen. Ob dann allerdings der Sinn des Satzes verstanden wurde, bleibt fraglich.

Insgesamt ist bei diesem Verfahren zu bezweifeln, ob das richtige Erraten eines Buchstabens tatsächlich auf die Konstruktion eines tieferen Verständnisses des Textinhalts schließen lässt. Der Satz im oben genannten Beispiel dürfte ohne Kontextualisierung und hinreichende Beschäftigung mit dem theologischen Hintergrund kaum verstanden werden.

Eine weitere Methode ist der Reaktionstest. Dieser Test ist quasi in den Text, der den Probanden gegeben wurde, direkt eingebaut, d.h. es sind Fragen oder Anweisungen eingefügt, die dem Probanden eine Handlung vorgeben. Hat der Proband diese richtig verstanden, reagiert er entsprechend. Teigeler hält fest, dass dieses Testverfahren bisher nie angewandt worden sei (vgl. Teigeler 1968, S. 54f), was auch über 40 Jahre später noch zutreffend zu sein scheint. Dieses Verfahren entspräche dem angedachten Vorgehen eines Tests mit Instruktionstexten (siehe unten).

Eine sehr verbreitetes Testverfahren ist der Multiple-Choice-Test (zu Deutsch: Mehrfach-Wahl-Aufgaben). Bei diesem wird über vorgegebene richtige und falsche Antworten versucht das Identifizieren/Wiederauffinden von Fakten (bei Vorliegen des Textes) oder Behalten von Fakten (bei nicht mehr Vorliegen des Textes) und Verstehen von Zusammenhängen zu erfassen. Die Schwierigkeiten dieses Verfahrens liegen in der Konstruktion der Fragen, bei denen sowohl die richtigen als auch die falschen Antworten gleich wahrscheinlich und plausibel sein müssen. Daneben werden meist viele Fragen für ein aussagekräftiges Ergebnis benötigt. Methodisch wird zudem noch der Einwand aufgeführt, dass dabei nicht nur die Textschwierigkeit eine Rolle spielen, sondern auch die Fragen selbst (vgl. Teigeler 1968, S. 54).

Abschließend stellt Teigeler das seiner Meinung nach beste Verfahren zur Ermittlung des Textverständnisses vor: der Lückenschließungstext (Cloze Procedure). Bei diesem wird einem Probanden ein Text gegeben bei dem jedes n-te Wort durch eine Lücke standardisierter Breite ersetzt wurde. Der Proband soll anschließend die fehlenden Wörter rekonstruieren, wobei die Anzahl der Treffer als Maß für die Schwierigkeit eines Textes gesehen wird (vgl. Teigeler 1968, S. 56f). Trotzdem, dass dieses Verfahren laut Teigeler angeblich das Beste sei, darf vom Autor angemerkt werden, dass hier ein Fehlschluss vorzuliegen scheint: von der Verständlichkeit eines Textes auf das Verstehen ebendieses zu schließen führt an der Sache vorbei. Mit dem Lückentest wird eher die sprachliche Kombinationsfähigkeit eines Probanden auf die Probe gestellt, aber kaum ermittelt, inwiefern er den Inhalt des Textes überhaupt wahrgenommen, geschweige denn verstanden hat, da insbesondere lokale Kohärenzbildung häufig auf Grund von grammatischen Analysen geschehen kann, ohne den globalen semantischen Gehalt des Textes zu berücksichtigen.

Von diesen vorgestellten Verfahren scheint sich insbesondere für die Messung von Lesen im Handlungskontext die Methode des Reaktionstests zu eignen. Dies kann von simplen Anweisungen wie „Heben Sie bitte den linken Arm.“ bis zu komplexen Aufgaben wie dem Bedienen eines Geräts an Hand einer schriftlichen Anleitung im Kontext einer annähernd realen Situation

mit etabliertem Textmaterial (vgl. Kapitel 0) geschehen. Auch bieten sich durch die Fortschritte der Computertechnologie neue mediale Möglichkeiten in Bezug auf bspw. interaktive Testprogramme, die in ihrer Komplexität vermutlich ähnlich variabel sein können. Warum dieses Format dennoch nicht gewählt werden kann wird im folgenden Kapitel deutlich, wenn die Rahmenbedingungen des Tests, der Test- sowie der Itementwicklung erörtert werden.

4.1.2. Rahmenbedingungen der Test- und Itementwicklung

Die Entwicklung des hier betrachteten Instruments erfolgte im Rahmen der ASCOT-Initiative. Daraus ergaben sich einige Rahmenbedingungen für dessen Konstruktion. Die ASCOT-Initiative des Bundesministeriums für Bildung und Forschung startete im Jahr 2011 mit dem Ziel, Instrumente für die Messung der am Arbeitsmarkt verwertbaren Handlungskompetenzen zu entwickeln, welche für die Ausübung einer qualitativ hochwertigen beruflichen Tätigkeit in einer sich wandelnden Arbeitswelt notwendig sind. Dabei steht die Bezeichnung ASCOT für die technologieorientierte Kompetenzmessung in der beruflichen Bildung. Zur Messung der entsprechenden Kompetenzen werden spezielle Instrumente und Methoden für verschiedene Berufsgruppen entwickelt⁸¹. Um die strukturellen Zusammenhänge zwischen beruflichen Kompetenzen und Basiskompetenzen zu erfassen, werden im Projekt MaK-adapt (Messung allgemeiner Kompetenzen – adaptiv) themenunabhängige Grundqualifikationen wie mathematische, naturwissenschaftliche und Lesekompetenzen fokussiert. MaK-adapt ist damit innerhalb der ASCOT-Initiative ein sogenanntes Querschnittsthema, welches die Aufgabe hat, Instrumente zur Erfassung von Erklärungsvariablen beruflicher Fachkompetenz für alle Vorhaben bereitzustellen.

Angelegt war das Projekt MaK-adapt von Dezember 2011 bis November 2014. Zunächst erfolgte die Festlegung auf ein Kompetenzmodell, danach eine Zusammenstellung des initialen Itempools, bspw. durch Recherche und Freigabe geeigneter vorhandener Itempools bzw. ausreichende Neukonstruktion von Items und Beschreibung der theoretischen Kompetenzmodelle in den jeweiligen Domänen (Mathematik, Naturwissenschaft oder in diesem Fall Lesen). Darauf wurden die vorhandenen und neukonstruierten Items computerisiert, d.h. so aufbereitet, dass eine Vorlage und Beantwortung der Fragen am PC möglich ist, und es erfolgte die Programmierung der verwendeten Fragebogensoftware, so dass zu Beginn des Winterhalbjahres die Planung und Durchführung der Kalibrierungsstudie zur Ermittlung der empirischen Schwierigkeiten der Items an Hand einer sequentielle Vorgabe der Items in einem Testheftdesign erfolgreich abgeschlossen wurde. Ziel der Kalibrierungsstudie war die Schätzung von Itemparametern für die anschließenden adaptiven Testungen. Daran anschließend erfolgte die Datenaufbereitung, Skalierung sowie verschiedene Analysen auf deren Basis eine Itemselektion für den späteren computerbasierten adaptiven⁸² Test (CAT) durchgeführt wurde. Nach der Erstellung einer vorläufige CAT-Version und verschiedener Simulationen erfolgte die Planung und Durchführung

⁸¹ vgl. www.ascot-vet.net

⁸² Zur Erläuterung siehe S. 180.

der Pilotierungsstudie im Sommerhalbjahr 2013. Die Pilotierungsstudie arbeitete mit dem produzierten adaptiven Algorithmus und hatte als Ziel die Erprobung und Optimierung ebendieses. Nach der darauffolgenden Datenaufbereitung und dem Cleaning wurden eine Optimierung der CAT-Versionen und die Erstellung der CAT-Endversionen vorgenommen. Ferner wurde ein Manual zu den CAT-Versionen erstellt und abschließend die CAT-Versionen als fertige adaptive Tests (inkl. Manual) den anderen Projekten zur Verfügung gestellt.

Auf Grund der Einbettung der Itementwicklung in diesen Kontext wurden jedoch spezielle Anforderungen an die Art und Weise der Messung der Lesekompetenz gestellt und vor allem bezüglich der Fragenkonstruktion. So sollte bspw. ein Instrument zur computerisierten, adaptiven Testung entwickelt werden. Daraus ergab sich eine gewisse Reihe von Bedingungen für die Entwicklung von Testitems. Diese waren:

Bezüglich der zeitlichen und technischen Rahmenbedingungen

- Das fertig pilotierte Messinstrument sollte schon nach 18 Monaten einsetzbar sein, so dass das Zeitbudget zur Neukonstruktion von Items sehr begrenzt war
- Die Konstruktion der Items erfolgte in MS PowerPoint, ohne dass Animationsmöglichkeiten genutzt werden konnten
- Die Dateigröße sollte möglichst kleingehalten werden, um eine akzeptable Übertragungsdauer unabhängig von der Datenleitung zu ermöglichen
- Es waren nur begrenzte Antwortformate (MC und kurze offene Antworten) möglich
- Die Testung erfolgte computerbasiert, d.h. Stimulus und Frage wurde den Probanden am PC präsentiert und wurden auch an diesem beantwortet.

Bezüglich der gestalterischen Anforderungen

- Es sollten möglichst kurze Stimuli und Fragen konstruiert werden, so dass viele unterschiedliche Fragen gestellt werden konnten
- Die Auswertung (richtig/falsch) der Antworten erfolgte durch ein Echt-Zeit-Computerprogramm unmittelbar nach der Eingabe der Antwort durch die Probanden
- Es sollte kein Testlet-Design⁸³ verwendet werden, da insbesondere beim adaptiven Testen eine stochastische Unabhängigkeit der jeweiligen Items notwendig ist. Sofern ein Stimulus dem Probanden in der Kalibrierung mehrfach vorgelegt wird, ist von einer anderen Itemschwierigkeit auszugehen, als wenn dieser Stimulus einmalig vorgelegt wird. Fraglich ist auch, inwiefern die Schwierigkeiten durch die Reihenfolge der Items

⁸³ ein Stimulus und dazu mehrere (theoretisch) unabhängige Fragen

mit gleichen Stimulus beeinflusst werden. Um diese Störungen (Positionseffekt) so gering wie möglich zu halten, wurde die Entscheidung getroffen auf ein Testlet-Design zu verzichten.

- Ferner konnte für den Bereich Lesen aus rechtlichen Gründen und auf Grund der Stichprobenorientierung nur wenige Items aus andern Studien adaptiert werden.

Die Entscheidung für ein Computergestütztes Adaptives Testverfahren (CAT) lässt sich wie folgt begründen⁸⁴: CAT ist eine spezielle Art der Testaufbereitung, bei der an Hand des Antwortverhaltens der Personen die Auswahl des folgenden Items bestimmt, so dass die Schwierigkeit der vorgelegten Items als maßgeschneidert auf die individuelle Leistungsfähigkeit abgestimmt, bezeichnet werden kann. Diese Art der Testdarbietung ist durch den Rückgriff auf Modelle der Item-Respons-Theorie (IRT) (siehe unten) möglich (vgl. Frey/Ehmke 2007).

Wie Frey/Ehmke (2007) ausführen ist durch adaptives Testen eine erhebliche Steigerungen der Messeffizienz möglich, ohne dass dabei Verluste an Messpräzision auf der Negativseite zu verzeichnen sind. Gleichzeitig kann im Vergleich zu herkömmlichen sequentiellen Tests eine höhere Differenzierungsfähigkeit über einen breiten Leistungsbereich erreicht werden (vgl. Frey/Ehmke 2007). Weiterhin können zusätzlich in motivationaler Hinsicht bezogen auf den Test insbesondere bei heterogenen Probandengruppen positive Effekte verzeichnet werden (vgl. Asseburg 2011). Anders als bei der konventionellen Aufgabendarbietung werden den Probanden beim adaptiven Testen nicht eine Vielzahl von Aufgaben vorgelegt, die entweder weit über oder weit unter der jeweiligen Leistungsfähigkeit angesiedelt sind. Konkret heißt dies in diesem Fall, dass abhängig vom Antwortverhalten der Probanden das nächste Item aus dem Itempool ausgewählt wird – bei einer falschen Antwort ein leichteres, bei einer richtigen ein schwereres (vgl. Abbildung 9)

⁸⁴ Vgl. im Folgenden auch bereits die Ausführungen in Ziegler et al. (2012).

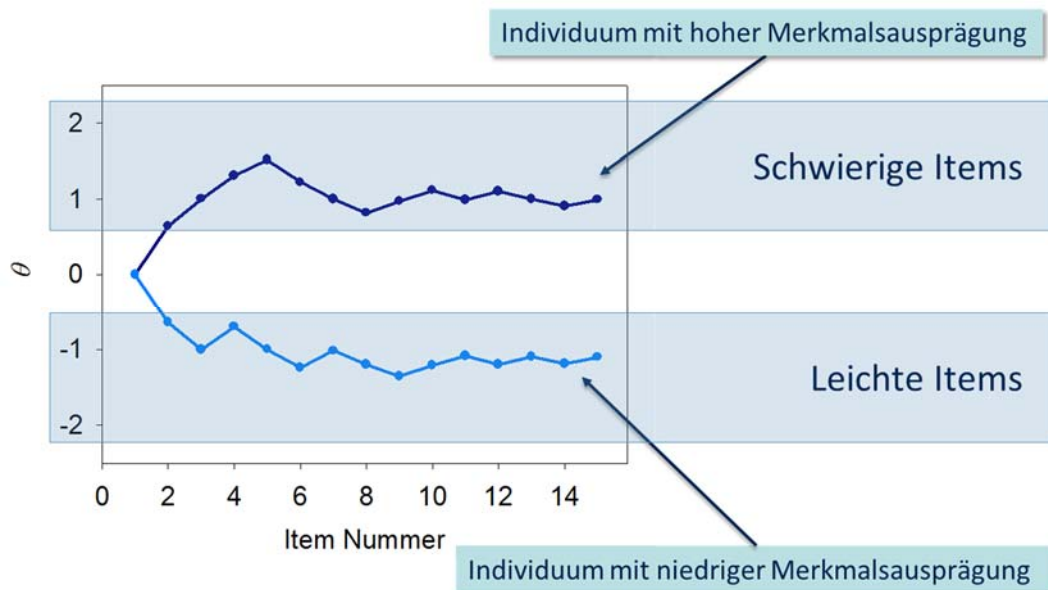


Abbildung 9 – adaptiver Testverlauf (Frey (2012))

Im Hinblick auf die anzunehmende Heterogenität von Absolventen substantiell unterschiedlicher Berufsfelder (kaufmännisch, medizinisch-pflegend, gewerblich-technisch) bietet das adaptive Testen somit ganz entscheidende Vorteile gegenüber konventioneller Testverfahren.

Bedacht werden muss jedoch auch, dass den erheblichen Vorteilen adaptiven Testens auch gewisse Nachteile gegenüberstehen. Zuallererst erfordert computerisiertes adaptives Testen einen hohen organisatorischen, finanziellen und technischen Aufwand, welcher sich jedoch nach der Einführung eines Testinstruments erheblich relativiert. Im ASCOT-Verbund erfolgt die Umsetzung mittels der kostenfrei zur Verfügung gestellten computerbasierten Testoberfläche MATE (vgl. Frey/Seitz 2010). Ein zurzeit noch ein evidenter Nachteil adaptiven Testens ist die Beschränkung exklusiv auf geschlossene Antwortformate. Das resultiert aus der Notwendigkeit, dass jede Antwort unmittelbar computerbasiert verarbeitet und bezüglich richtig oder falsch an Hand vorgegebener Lösungen interpretiert werden muss, so dass davon abhängig die Auswahl und Zuweisung der weiteren Aufgabe erfolgen kann. Offene Antwortformate und interaktive Textfeatures können daher bisher noch nicht verwendet werden (vgl. Frey/Ehmke 2007).

Da also das Programm MATE derzeit ausschließlich dichotome Items im Einfach-Antwort-Auswahl-Format erlaubt, stellen sich hohe Anforderungen an die Konstruktvalidität der Items. Unter anderem muss besondere Augenmerk darauf gerichtet werden, dass anstatt der Lesefähigkeiten nicht vor allem Vorwissen und verbale Intelligenzleistungen gemessen werden (vgl. Jordan 2011). Die Items sollten einerseits zugleich authentisch sein, dennoch auch derart entfremdet, dass Informationen zu ihrer Bearbeitung tatsächlich den vorgelegten Stimulus entnommen werden müssen und nicht schon auf Grund der Vertrautheit mit den Sachverhalten (also ohne zu lesen) gelöst werden können.

Diese Anforderungen boten wie dargestellt einige Herausforderungen für die Itemkonstruktion. Zeitliche Beschränkungen bei der Entwicklung lassen oftmals nicht den Raum für alle angestrebten Voranalysen und Berücksichtigung von Parametern, die anzustreben wären. Die Neuentwicklung vieler Items ist zeitintensiv und Ressourcen beanspruchend, da nicht auf bereits etablierte Materialien zurückgegriffen werden kann, welche bereits gründlich bezüglich verschiedener Faktoren analysiert wurden.

Die computerisierte Darstellung von Lesematerial kann kaum ohne die Berücksichtigung medialer Effekte erfolgen, um somit zunächst den Einfluss des Testmediums zu untersuchen, wurde zusätzlich in der Pilotierungsphase, deren Kernziel die Testung des entwickelten adaptiven Algorithmus war, eine Vergleichsuntersuchung zwischen computerbasierter und papierbasierter Testung mit einigen Items durchgeführt (die diesbezüglichen Ergebnisse und Erkenntnisse sind in Kapitel 0 zu finden).

Die Notwendigkeit einer sofortigen Auswertbarkeit lässt trotz guter Gründe kaum ein anderes Testverfahren als Single-Choice zu. Bei kurzen, wenige (1-2) Wörter umfassenden offenen Antworten müssen für offenen Antwortformate alle als richtig akzeptierten Möglichkeiten (Schreibweisen, Darstellungsmöglichkeiten,...) im Programm hinterlegt sein. Bei der Konstruktion besonders ins Gewicht fiel, dass in MS PowerPoint eine angemessen große Schriftgröße gewählt werden muss, um die Lesbarkeit zu gewährleisten – dadurch erhöht sich jedoch rasch die Seiten- (bzw. Folien-)zahl. Bezogen insbesondere auf die Darstellung von depiktionalen Strukturen ist die Balance zwischen notwendiger Genauigkeit, d.h. Größe, Detailstufe (Auflösung) und Farbe auf der einen Seite und geforderte geringer Dateigröße auf der anderen Seite auszuloten.

Dennoch wurde versucht die in Kapitel 3.1.5 dargestellten Parameter bezüglich der Verständlichkeit von Texten als schwierigkeitsbeeinflussende Parameter zu berücksichtigen. Bevor die Beschreibung der eigentlichen Itemkonstruktion erfolgt, wird im Folgenden erläutert, welche konkreten theoretischen Annahmen dem Kompetenzmodell zugrunde liegen und wie sich dieses gestaltet.

4.1.3. Theoretisches Kompetenzmodell⁸⁵

In den aktuellen LSA-Studien der OECD (PISA, PIAAC) wurden und werden im Hinblick auf die Gestaltung der Leseanforderungen in Testitems drei Aspekte des Lesens unterschieden: „Identifizieren und Extrahieren“, „Integrieren und Interpretieren“ sowie „Reflektieren und Bewerten“ (vgl. Artelt et al. 2001; Naumann et al. 2010). Werden diese Aspekte in Bezug auf textbezogene Lesefunktionen und deren ablaufende Vorgänge betrachtet, können sie auch als Unterschiede auf Seiten der kognitiven Verarbeitungsprozesse verstanden werden.

⁸⁵ Vgl. im Folgenden auch bereits die Ausführungen in Ziegler et al. (2012).

Aufgaben im Bereich von Identifizieren und Extrahieren sowie im Bereich von Integrieren und Interpretieren erfordern „lediglich“ textimmanente Verstehensleistungen. Es kann davon ausgegangen werden, dass die erforderlichen Informationen weitgehend aus den zur Verfügung stehenden Repräsentationsformaten entnommen werden können. Reflektieren und Bewerten dagegen erfordert größtenteils wissensbasierte Verstehensleistungen. In diesem Bereich wird an den Lesenden die Anforderung gestellt, dass er den Text mit eigenen Wissensbeständen, Erfahrungen und Ideen in Beziehung setzt (vgl. Artelt et al. 2001; Naumann et al. 2010).

Die oben angeführten „Aspekte des Lesens“ können noch als sehr stark an den Gegenständen schulischen Lernens orientiert bezeichnet werden. Für die Lösung von Aufgaben im Bereich des Reflektierens und Bewertens ist häufig der Aufbau einer mentale Repräsentationen der Kommunikationsabsicht der schriftlichen Dokumente (Kommunikationsebene) oder des Genres (Genreebene) erforderlich (vgl. Schnotz/Dutke 2004). Konträr erfordern die Aufgaben des Integrierens und Interpretierens eher ein textanalytisches Vorgehen, indem grundlegende Beziehungsstrukturen zwischen (verschiedenen) Textteilen – bspw. rhetorische Relationen – erfasst werden. Dies setzt primär die Repräsentation der Textbasis voraus (vgl. OECD 2009). Die Subskala Reflektieren und Bewerten ist demnach gegenüber den anderen Skalen komplexer – was sich auch darin bestätigt, dass die Fähigkeiten der Schüler in Deutschland auf dieser Skale signifikant unter den Leistungen auf den beiden anderen Skalen liegen (vgl. Naumann et al. 2010).

Die Aufgaben des Bereichs Identifizieren von Informationen werden von der OECD (2009) auch als Nutzungsaufgaben bezeichnet, wobei aber wahrscheinlich darauf bezogen wird, dass der Text überwiegend verwendet wird, um Informationen zu gewinnen – bspw. einer Anzeige den Beginn eines Filmes im Kino oder einem Kuchenrezept die Mengenangabe für Mehl zu entnehmen. Die Schwierigkeit dieser Art von Aufgaben kann überwiegend dadurch modifiziert werden, dass die Informationen an gewöhnlichen oder ungewohnten Stellen im Text stehen (vgl. Naumann et al. 2010). Nutzung wird in diesem Kontext also eher verstanden als die Nutzung eines Textes zum Zweck der Informationssuche und weniger unter der Perspektive, dass die Informationen unmittelbar zur Ausführung einer Handlung „genutzt“ werden.

Wie jedoch bereits ausführlich dargelegt scheint im beruflichen Kontext aber genau dieser Aspekt der „Nutzung von Informationen“ relevant: sehr häufig wird mit dem Ziel gelesen, die Informationen so zu erschließen, dass unmittelbar in eine Handlung umgesetzt werden können. Das obige Beispiel weiter gedacht bedeutet, dass einem Kuchenrezept nicht nur die Menge an Mehl entnommen werden muss, sondern auch die im Rezept enthaltenen Anweisungen zur Erstellung des Kuchens umgesetzt werden müssen (Ziegler et al. 2012). Es ist wahrscheinlich, dass für die Aufgabe die Mengenangabe des Mehls zu entnehmen eine Oberflächenrepräsentation des Textes ausreichen dürfte, während die Ausführung des Rezepts wahrscheinlich den Aufbau eines mentalen Modells erfordert (vgl. Schnotz et al. 2011). Dies ist für einen Konditor aller Wahrscheinlichkeit nach leicht, insbesondere sofern er bereits einiges an Erfahrung im Backen von Kuchen gesammelt hat und so bereits auf ein mentales Modells des Kuchenbackens (welches in diesem Fall bspw. nicht text-, sondern erfahrungsbasiert ist) zurückgreifen kann, in

das die textbasierten, d.h. möglicherweise neuen oder zusätzlichen Information aus dem jeweiligen Rezept integriert werden müssen. Handelt es sich jedoch um einen Laien, der seinen ersten Kuchenbacken will, geht die Leseanforderung der gleichen Situation weit über die bloße Identifikation von Informationen hinaus, da Handlungsprozesse an Hand einer Textbasis rekonstruiert werden müssen. Trotzdem handelt es sich um eine textimmanente Verstehensleistung.

Daher scheint es unzulänglich, die Leseanforderungen nur entsprechend der genannten Aspekte des Lesens auszuarbeiten. Für die Erfassung funktionaler Lesefähigkeiten sind sie wenig kompatibel. Als erste Annäherung wurde zur Modellierung der Leseanforderungen auf den Entwurf/das Modell von Schnotz/Dutke (2004) das „Modell des integrativen Text- und Bildverstehens“ zurückgegriffen (vgl. Kapitel 0), um funktionale Leseanforderungen zu bestimmen und zu ordnen. Für diese Überlegung sprechen auch empirische Befunde, die zeigen, dass dimensionale Unterschiede in den Fähigkeitsausprägungen sich vor allem hinsichtlich kognitiver Anforderungsparameter zeigen (vgl. Husfeldt 2001; Schnotz et al. 2010). In seiner hier zentralen Form sieht das zugrundeliegende Modell der kognitiven Anforderungen folgendermaßen aus:

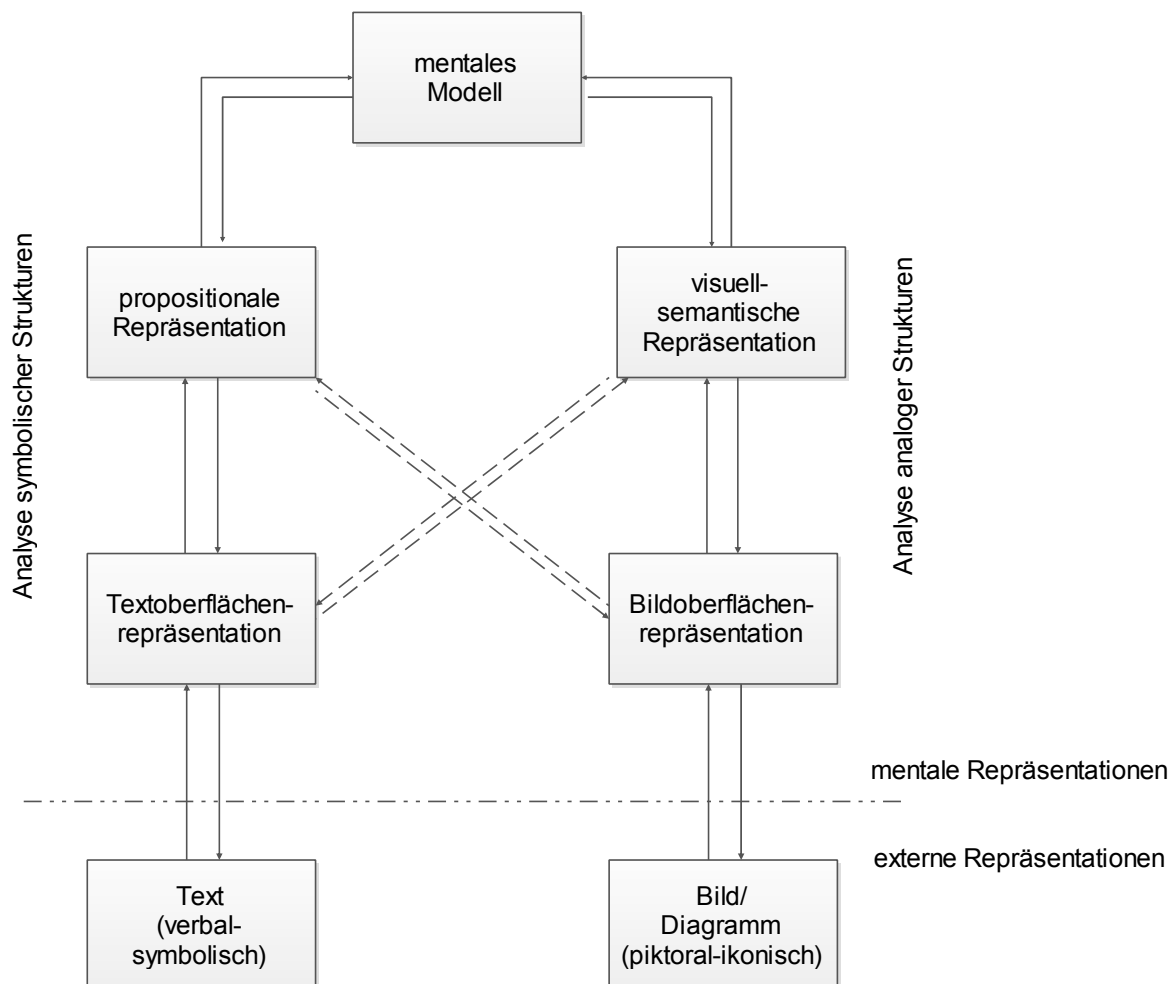


Abbildung 10 – Lesen in beruflichen Handlungskontexten (Modellierung angelehnt an Schnotz/Bannert bzw. Schnotz/Dutke)

Das Modell in seiner hier verwendeten Form (Abbildung 10) umfasst, wie bereits dargestellt, einen deskriptionalen und ein depiktionalen Repräsentationszweig mit den entsprechenden

mentalenen Repräsentationen, die sich hinsichtlich Beständigkeit und möglicherweise auch Verstehensqualität unterscheiden.

Die Verarbeitungsprozesse beim Bild- bzw. Diagrammverstehen verlaufen im Unterschied zum Textverstehen zunächst primär über eine Strukturabbildung (d.h. analoge Verarbeitung) ausgehend von der perzeptuellen Repräsentation zum mentalen Modell. Beim Textverstehen werden symbolische Verarbeitungsprozesse angenommen, welche von der Textoberfläche ausgehen und über eine propositionale Repräsentation der Textbasis verlaufen (vgl. Schnotz/Dutke 2004).

Im Gegensatz zu Schnotz wird hier von einer Zwischenstufe zwischen der Bildoberflächenrepräsentation und dem mentalen Modell ausgegangen. Angenommen werden kann, dass diese weitere kognitive Verarbeitungsebene insbesondere bei wenig analogen, jedoch nichtsdestotrotz graphischen Strukturen wie bspw. logischen Bildern zum Tragen kommt. Möglicherweise kann diese Stufe bei realistischen Bildern (vgl. Kap. 0) übersprungen werden. Weitere Forschung insbesondere zur Analyse und zum Verständnis logisch-abstrakter Bilder scheint lohnend.

Nach Schnotz (2005) verbleiben die Repräsentationen der Textoberfläche bzw. visuelle Wahrnehmungen eines Diagramms bzw. Bildes im Arbeitsgedächtnis und sind eher flüchtig. Dadurch wird lediglich die Widergabe von Text- und Bilddetails ermöglicht, jedoch kein Verstehen im engeren Sinne von im Text und Bild enthaltenen Bedeutungen. Propositionale Repräsentationen sind sodann das Ergebnis einer semantischen Analyse von Text oder Bild/Diagramm. Die propositionale Repräsentation eines Textes ist beständiger als die Textoberfläche, besitzt jedoch gegenüber dem mentalen Modell eine höhere Vergessensrate (vgl. Schmalhofer/Glavanov 1986). Ferner ermöglichen propositionale Textrepräsentationen ein Verstehen des Textinhaltes, jedoch noch keine von der eigentlichen Textbasis gelöste Vorstellung zum Sachverhalt (vgl. Schnotz 2006a, 2006b). Propositionale Repräsentationen auf der Grundlage depiktionaler Darstellungen stellen dagegen wahrscheinlich das Ergebnis einer konzeptgeleiteten Analyse dar, welche parallel zur mentalen Modellkonstruktion erfolgt. Diesem mentalen Modell werden über Ablese- und Inspektionsprozesse relevante Informationen entnommen und in Form propositionaler mentaler Repräsentationen gespeichert (vgl. Schnotz/Dutke 2004, S. 75).

Die zum Verständnis eines Textes notwendigen Verarbeitungsprozesse sind insgesamt in hohem Maße vom jeweiligen Vorwissen des Lesers abhängig. Die Annahme, dass die mentale Modellbildung eines im Text repräsentierten Sachverhalts in jedem Fall eine propositionale Repräsentation der Textbasis erfordere, wird heute nicht mehr in ihrer absoluten Form aufrechterhalten (vgl. Zwaan 2004; Zwaan/Madden 2005; Leopold 2009). Wahrscheinlicher als derart lineare Annahmen sind parallel und auch quer verlaufende Prozesse (entsprechend sind auch die gestrichelten Diagonalen im Modell zu interpretieren). So ist z.B. naheliegend, dass bei bekannten Sachverhalten schon ausgehend von der Textoberfläche mentale Modelle aktiviert werden, gleichfalls ist anzunehmen, dass beim Lesen vertrauter depiktionaler Darstellungen parallel propositionale Repräsentationen gebildet werden.

Analog verlaufen die rezeptiven Verarbeitungsprozesse nicht singulär in eine Richtung (bspw. ausgehend von der externen Repräsentation zu multiplen mentalen Repräsentationen). Stattdessen können die aus einem Bild bzw. Diagramm entnommenen und mental modellierten Informationen in eine externe symbolische Repräsentation transformiert werden, in der der Sachverhalt verbalisiert wird und vice versa. Basierend auf diesen im Modell entworfenen Prozessen sowie in Anlehnung an Kirsch (1999) sollen mit dem Instrument drei grundlegende Anforderungsklassen abgebildet werden: Identifizieren, Integrieren und Generieren.

Abbilden lässt sich dies in etwa wie folgt:

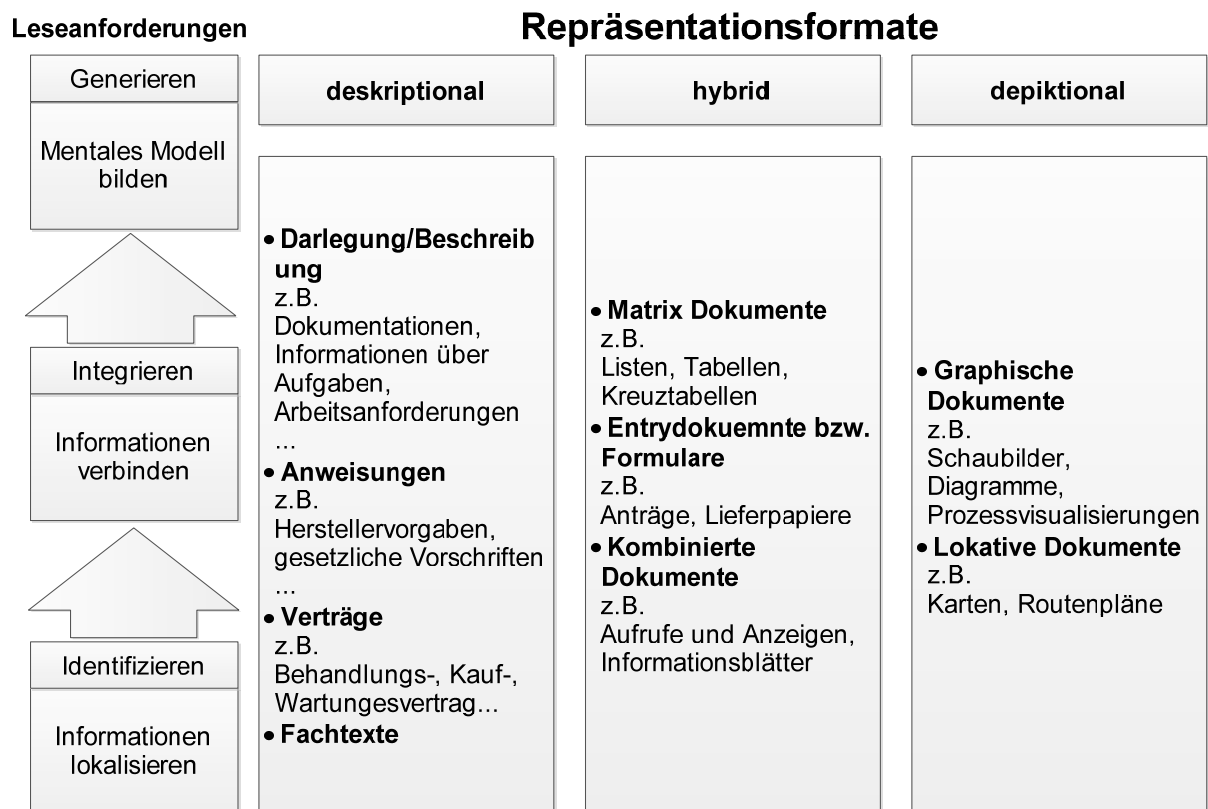


Abbildung 11 – Repräsentationsformate und Anforderungsklassen⁸⁶

Zur Spezifikation: Der Vorgang des Identifizierens verlangt vom Leser lediglich das (Wieder-) Erkennen von Bild- oder Textelementen. Die dafür notwendigen Informationen dürften bereits auf Grundlage einer Oberflächenrepräsentation des Textes bzw. der visuellen Wahrnehmung eines Bildes zur Verfügung stehen. Die Anforderungshöhe kann über die Lösungsparameter der Aufgabe modifiziert werden. Es ist anzunehmen, dass beispielsweise eine Information in einem Text leichter zu identifizieren ist, sofern diese auch in der Fragestellung explizit benannt wird (gleiche Begriffe auffinden). Soll dagegen ein zur Fragestellung synonyme Begriff im Text identifiziert werden, steigt damit der Anforderungsgehalt. Gleichartig sind vermutlich auch die Anforderungen, wenn die Entsprechung eines verbalen Begriffs in einem Bild oder Diagramm

⁸⁶ Die Terminologie orientiert sich hier nicht primär an Kirschs (1999) Arbeiten, sondern versucht eine Brücke zu Schnotz (2005) Theorien und Analysen zu schlagen. Daher bezieht sich die Benennung der Repräsentationsformate auch auf die von Schnotz verwendete Terminologie zuzüglich der Kategorie der hybriden Formate.

identifiziert werden soll. Weiterhin kann durch die Platzierung von irrelevanten Informationen im Dokument der Anforderungsgehalt beim Identifizieren weiter gesteigert werden (vgl. Kirsch 1999).

Bezüglich des Integrierens von Informationen lautet dagegen die Annahme, dass mindestens eine propositionale Repräsentation erforderlich ist. Diesbezüglich kann „Integrieren“ heißen, mindestens zwei Informationen, welche innerhalb eines Textes verstreut sind, aufeinander zu beziehen und deren Bedeutungen integrativ zu verarbeiten. Analog sind die Anforderungen, wenn mehrere Informationen innerhalb einer depiktionalen Repräsentation integrativ zu einer Bedeutung verarbeitet bzw. einfache Relationen zwischen Bildinformationen hergestellt werden müssen. Integrieren wird zudem gefordert, wenn zwischen Text- und Bildinformationen Beziehungen hergestellt werden müssen (bspw. Diagramm und Legende). Eine Variation der Schwierigkeit in den Integrationsaufgaben kann durch die Anzahl der zu integrierenden Informationen, durch die Stärke ihrer Verbindung und ebenfalls durch die Platzierung der Informationen sowie durch Informationsumfang und -dichte des Dokuments erfolgen (vgl. Kirsch 1999).

Aufgaben im Bereich des Generierens sollen so beschaffen sein, dass zur Bearbeitung eine mentale Modellbildung erforderlich wird. Erreicht werden soll dies bei der Aufgabenstellung über einen Wechsel des Repräsentationsformates (Text < > Bild). In deskriptionalen Repräsentationen dargebotene Sachverhalte müssen dementsprechend in depiktionale Formate transformiert werden, während solche Bild- bzw. Diagramme bzw. deren Interpretation verbalisiert werden müssen. Hier kann eine Variation der Schwierigkeit über die Komplexität des mentalen Modells erfolgen, welches ebenfalls vom Informationsumfang und der -dichte des externen Repräsentationsformats abhängen dürfte.

Eine weitere Dimension sind die unterschiedlichen Repräsentationsformate, die – wie in Kapitel 0 bzw. 0 ausgeführt – unterschieden werden in deskriptionale Formate (Fließtexte), depiktionale Formate (Bilder, Graphiken, Charts, etc.) und hybride Formate, die Elemente von beiden Kategorien enthalten.

Auf Grund der bereits in Kapitel 0 ausgeführten strukturellen Unterschiede der verschiedenen Repräsentationsformate bezüglich ihrer repräsentativen Stärke war es nicht immer möglich, den vollständigen Informationsgehalt eines Textes zu visualisieren. Die externe Visualisierung soll somit also lediglich als ein Indikator dafür betrachtet werden, dass im Text repräsentierte Sachverhalte mental modelliert wurden.

Weiterhin soll der zentralen Funktion des Lesens im beruflichen Handlungskontext, der Handlungsausführung, durch die Aufgabenstellung Rechnung getragen werden. Dies bedeutet, dass das Leseziel, Informationen zu finden, um eine berufliche Handlung auszuführen, deutlich in der Aufgabenstellung betont wird.

In Kapitel 3.1.5 wurde versucht an Hand der Textrezeptionsforschung herauszufinden, welche „Stellschrauben“ in einem Item möglicherweise angepasst werden können, um seine Beantwortung wahrscheinlicher, d.h. es selbst leichter zu machen und vice versa. Weitere Hinweise auf Variationsmöglichkeiten der Schwierigkeit wurden bereits in diesem Kapitel genannt. An diesen wurde sich während der Konstruktion der Items orientiert. In Kapitel 4.2.4 folgt eine differenzierte Analyse der Items unter Aspekten der Textverständlichkeitsforschung.

4.1.4. Umsetzung

Wie bereits ausgeführt, wurde insbesondere der berufliche Kontext auf Itemebene integriert durch die Vorgabe des Leseziels/der Leseintention einen Text zu rezipieren, um eine andersgeartete Aufgabe zu erfüllen, bspw. „Lesen Sie bitte, ..., um ... zu...“ umzusetzen. Es wurden die speziellen Text(misch)formate (Hybride), die im beruflichen Bereich auftreten, besonders berücksichtigt und versucht in eine berufsbezogene Lesesituation zu etablieren. Die mentale Modellierung der Handlung, die im Anwendungsbezug durch die zentrale Funktion des Umsetzens notwendig erscheint, wurde durch die Anforderung einen Moduswechsel zur Lösung der Aufgabe vollziehen zu müssen, bei vielen Aufgaben versucht adäquat abzubilden.

Bei der Entwicklung der Testfragen, war das Ziel ein gewisses Spektrum an antizipierten kognitiven Anforderungsbereichen mit ansteigenden Schwierigkeiten (identifizieren, integrieren, generieren) und drei Oberkategorien an Repräsentationsformaten (deskriptional, hybrid, depiktional) abzudecken (vgl. Kapitel 4.1.3). Weiterhin sollte insbesondere der Bereich der hybriden Items einen größeren Anteil haben. Insgesamt konnten in der zur Verfügung stehenden Zeit 73 Items konstruiert werden.

Diese Items lassen sich nach dieser ersten Phase in folgender Matrix analog zu (Abbildung 11) wiedergeben. In den einzelnen Zellen stehen die bei der Konstruktion fortlaufend vergebenen Nummern, sogenannte Item IDs.⁸⁷

⁸⁷ Zum Verständnis: Werden im Laufe der Untersuchung Items gelöscht, bleiben die IDs stabil, d.h. die Nummerierungen werden nicht aufgerückt.

Tabelle 5 - Itempool (gesamt) Matrixdarstellung

		Repräsentationsformat			
		deskriptional	hybrid	depiktional	gesamt
Anforderungsbereich	identifizieren	06		08	23
		25	02	65	
		32	15 48	20 67	
		37	30 56	45 68	
		49	39	59 72	
		51		60 73	
	integrieren		03		20
		29	04 43	10 23	
		50	18 46	12 33	
		66	34 52	14 53	
		70	36 58	17	
	generieren	05	07 26		30
		16 40	09 31	01	
		27 55	11 41	24 62	
		28 57	13 42	35 63	
		38 71	19 44	54 69	
			21 47		
			22 64		
	gesamt	19	29	25	73

Es konnten alle Bereiche solide abgedeckt werden. Die mittlere Repräsentationskategorie weist eine leichte Schwerpunktsetzung auf (39% der Testfragen liegen im hybriden Format vor). Dies wird unter dem Gesichtspunkt der Authentizität positiv gewertet, da derartige hybride Formate im Berufsalltag überwiegen (vgl. Kapitel 2.1.2).

Der oben angesprochene Moduswechsel, welcher als zentral für das Kompetenzmodell bezeichnet werden kann, kann sowohl von der deskriptionalen Repräsentation in eine depiktionale erfolgen, als auch vice versa. Zur Illustration werden unten verschiedene Items bezüglich dieser Möglichkeiten vorgestellt, an Hand derer werden auch die anderen Kriterien der Entwicklung kurz illustriert.

Einfache Items, auf der kognitiven Dimension des *Identifizierens*, können depiktionaler oder deskriptionaler Natur sein. Auch können trotz des vermeintlich einfachen Anforderungsbereichs innerhalb einer Klasse noch Differenzierungen bezüglich der Schwierigkeit stattfinden wie im Folgenden gezeigt wird.

Zunächst ein depiktionales Item, welches auch als Ableseaufgabe bezeichnet werden kann:

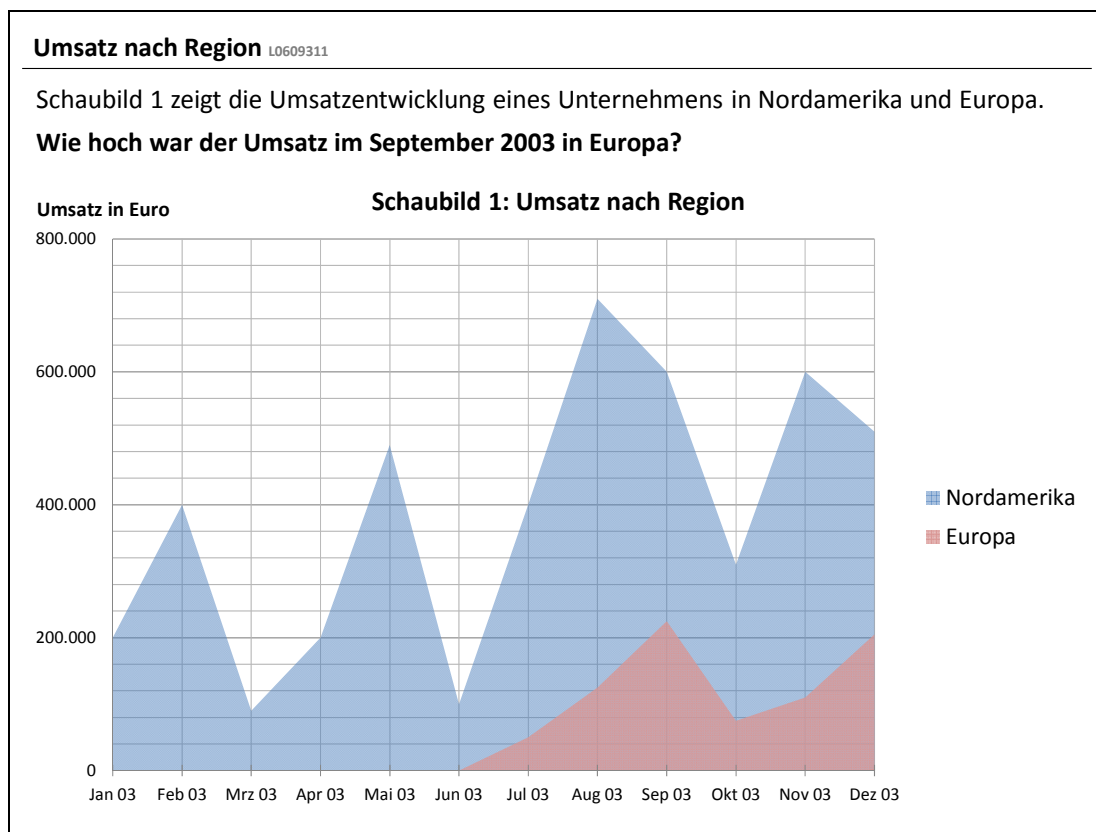


Abbildung 12 – Item 60 Seite 1

Wie hoch war der Umsatz im September 2003 in Europa?

Bitte kreuzen Sie an.

- ☐ ca. 220.000 €
- ☐ ca. 600.000 €
- ☐ ca. 200.000 €
- ☐ ca. 80.000 €

Abbildung 13 – Item 60 Seite 2

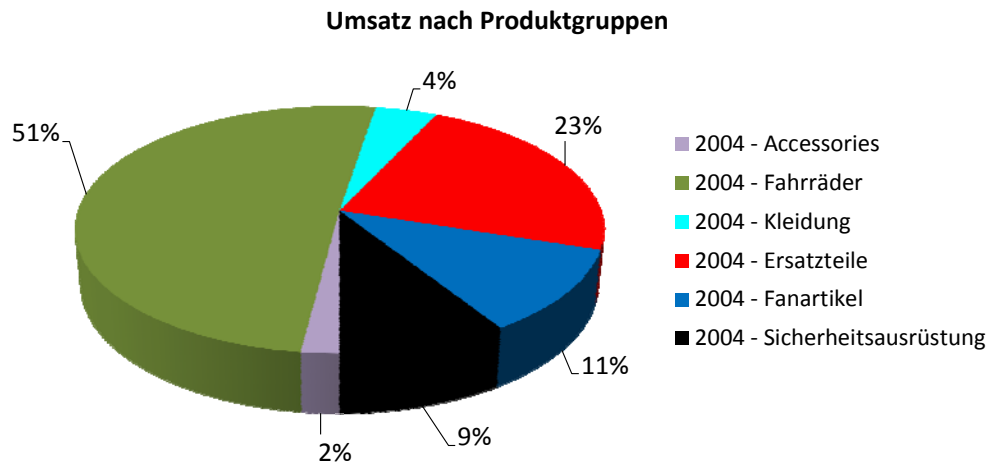
Gefragt wird nach dem Umsatz im September 2003 in Europa. Auch wenn es global als verhältnismäßig einfach bezeichnet werden kann, muss der Proband eine Reihe von kognitiven Aufgaben erfüllen: zunächst muss im Diagramm erkannt werden, dass sowohl die Umsätze von Nordamerika als auch von Europa durch ein Flächendiagramm (statt wie üblicher wäre durch ein Linien- oder Balkendiagramm) dargestellt werden. Die Monate sind abgekürzt, die Kürzel können aber als bekannt vorausgesetzt werden. Die anspruchsvollste Leistung ist die richtige Interpretation der unbeschrifteten Zwischenlinien. Dazu muss erkannt werden, dass die Linien keine Abstände von 1.000, 10.000 oder 20.000 markieren, sondern größer sind. Theoretisch ermittelt der Proband an Hand der Darstellung eine Zahl von mindestens 200.000 (exakt: 220.000). Bei der Betrachtung der Distraktoren scheidet 80.000 (fehlerhaftes Ablesen) und 600.000 (Betrag für Nordamerika) aus. Die Entscheidung zwischen 200.000 und 220.000 muss auf Grund der oben dargestellten Tatsache getroffen werden, dass die Zwischenlinien größere Abstände darstellen und die Fläche sodann die Markierung „200.000“ überschreitet und somit die Antwort „220.000“ richtig sein muss.

Damit lässt sich festhalten, dass auch ein Item des Bereichs Identifizieren verhältnismäßig komplex und anspruchsvoll sein kann. Im Vergleich dazu ein auf die kognitiven Operationen (vgl. Parameter aus Kapitel 0) bezogen als einfacher zu klassifizierendes Item ebenfalls depiktionaler Einordnung ist das Folgende:

Umsatz nach Produktgruppen L0599311

Das folgende Diagramm zeigt den Umsatz nach Produktgruppen im Jahr 2004.

Bitte finden Sie heraus, welches die schwächste Produktgruppe ist.



Bitte tragen Sie hier ein, welches die schwächste Produktgruppe ist:

Abbildung 14 – Item 59

Hier muss lediglich erkannt werden, welches die kleinste %-Zahl und/oder die kleinste Fläche ist und die gemäß der Legende zuzuordnende Produktgruppe in das Antwortfeld eingetragen werden (Rechtschreibfehler wurden, sofern sie klar erkennen ließen, welche Gruppe gemeint ist, ignoriert).

Ein analog ähnlich einfach konstruiertes Item aus dem Bereich *Identifizieren* jedoch deskriptiven Formats ist dieses:

Bitte lesen Sie den Text und finden Sie heraus, warum Schwimmen im Teich für Enten weniger gefährlich ist als für Menschen.

In der **Wochenzeitung** DIE ZEIT gibt es regelmäßig eine Antwort auf eine Frage von Leserinnen und Lesern.

Blitztod im Teich

Man sollte nicht im See schwimmen, wenn es gewittert, weil die Gefahr besteht, dass man dabei vom Blitzschlag getroffen wird. Meine Tochter (8 Jahre) wandte aber ein, dass den Enten und den Fischen doch wohl auch nichts passiert. Ist es nun gefährlich, als Mensch bei Gewitter im See zu schwimmen? *Ralf J. Radinski, Berlin*

Der Ratschlag, bei Gewitter nicht im See zu baden, ist wohlbegründet. Denn auf der großen Seeoberfläche ist ein einsamer menschlicher Schwimmer der höchste Punkt weit und breit, und bekanntlich sucht der Blitz sich solche Punkte mit Vorliebe beim Einschlagen aus.

Was ist nun mit den Tieren, nach denen Ihre Tochter fragt? Da gilt zunächst einmal: Je kleiner, desto ungefährlicher – in eine Ente schlägt der Blitz nicht so schnell ein wie in einen Menschen, und in kleinen Körpern entstehen in einem elektrischen Feld nicht so hohe Spannungen wie in einem großen. Und was die Fische angeht: *Unter* Wasser ist man erheblich sicherer als *auf* dem Wasser, weil man ja nicht mehr ein bevorzugtes Ziel für den Blitz ist. Und weil das Wasser ein guter Leiter ist, verteilt sich die Energie des Blitzes in alle Richtungen, sodass Wasserbewohner und Taucher schon einige Meter Entfernung vom Blitzschlag wenig zu befürchten haben – je tiefer, desto sicherer.

Deshalb schwimmen nach einem Gewitter nicht Tausende von toten Fischen auf dem See, sondern allenfalls ein paar Exemplare, die sich nahe an der Oberfläche und nahe am Einschlagsort des Blitzes befanden. *Christoph Drösser*

Drösser, Christoph: „Stimmts? Blitztod im Teich In: Die Zeit, 12.06.2003,Nr. 25

Abbildung 15 – Item 6 Seite 1

Warum schlägt der Blitz nicht so schnell in eine Ente wie in einen Menschen ein?

Bitte kreuzen Sie an. Weil die Ente ...

- ☐ ... bei Gewitter unter Wasser taucht.
- ☐ ... sich auf der Wasseroberfläche befindet.
- ☐ ... schnell ans Ufer schwimmt.
- ☐ ... sehr klein ist.

Abbildung 16 – Item 6 Seite 2

Bei diesem Item⁸⁸ geht es um die Frage, warum eine Ente im Wasser nicht so wahrscheinlich vom Blitz getroffen wird wie ein Mensch. Mit dieser Information liest der Proband den präsentierten Text (Quelle: Drösser 12.06.2003) vermutlich sehr zielgerichtet, fokussiert auf diesbezügliche Informationen (selektives Lesen, bestimmt von der gestellten Anforderung und der Intention diese zu erfüllen). Kommt er zu der Phrase „je kleiner desto besser...“ und die weiteren Erläuterungen hat er seine Antwort gefunden. Betrachtet er sodann die Distraktoren finden sich zwar die meisten in irgendeiner Form im Text wieder, jedoch hat nur einer einen Bezug zur Größe der Ente – somit kommt der Proband relativ schnell und unkompliziert zu der richtigen Antwort.

Abschließend ein vermutlich schwierigeres Item, welches aber auch noch als Identifikationsleistung klassifiziert wurde – ebenfalls deskriptionalen Repräsentationsformats ist folgendes:

Sich auf ein Einstellungsgespräch vorbereiten L0499111

Sie bereiten sich auf ein Einstellungsgespräch vor und wollen sich über den Arbeitgeber informieren. **Lesen Sie die Beschreibung aufmerksam durch, um die geeigneten Informationsquellen zu finden!**

Einstellungsgespräche

Vorbereitung
Versuchen Sie, mehr über den Arbeitgeber zu erfahren. Welche Produkte stellt das Unternehmen her, welche Methoden oder Verfahren wendet es an? Diese Informationen können Sie auf der Internetseite des Unternehmens oder bei der Handelskammer einholen. Finden Sie mehr über die Arbeitsstelle heraus. Würden Sie jemanden ersetzen, oder ist es eine neu eingerichtete Position? In welchen Abteilungen oder Geschäften würden Sie arbeiten? Informationen darüber können Sie bei der Bundesagentur für Arbeit (BA) einholen. Analysieren Sie die Stellenanzeige. Welche Erwartungen hat das Unternehmen an potentielle Stelleninhaber? Beschreibungen verschiedener Berufe und Arbeitsgebiete sind meist im Internet erhältlich.

Das eigentliche Gespräch
Stellen Sie Fragen über die Arbeitsstelle und den Arbeitgeber. Beantworten Sie alle Ihnen gestellten Fragen klar und deutlich. Nehmen Sie etwas zum Schreiben und Ihre Bewerbungsunterlagen mit.

Nach dem Gespräch
Notieren Sie sich die Kernpunkte des Gesprächs. Vergleichen Sie die Fragen, die Ihnen Schwierigkeiten bereitet haben mit denen, die es Ihnen ermöglichten, Ihre starken Seiten herauszustellen. So ein Rückblick wird Ihnen helfen, sich auf zukünftige Einstellungsgespräche vorzubereiten. Wenn Sie wollen, können Sie darüber mit dem Stellenvermittler oder dem Berufsberater Ihres örtlichen Arbeitsamtes sprechen.

Abbildung 17 – Item 49 Seite 1

⁸⁸ Quelle: Sächsisches Bildungsinstitut: Kompetenztest Klassenstufe 8 Hauptschulbildungsgang Schuljahr 2010/2011 Fach Deutsch bzw. IQB-Item aus VERA 8, leicht modifiziert (Leseauftrag und Frage)

Welche Informationen finden Sie nach dieser Beschreibung bei der BA?

Bitte kreuzen Sie an.

- ☐ Informationen über Produkte, die das Unternehmen herstellt.
- ☐ Informationen zu erwarteten Bewerberzahlen.
- ☐ Informationen, ob die Position neu eingerichtet wurde.
- ☐ Informationen über mögliche Methoden und Verfahren.

Abbildung 18 – Item 49 Seite 2

Hier muss der Proband verhältnismäßig viele mentale Schritte vollziehen. Er muss zu allererst verstehen, dass es zunächst nur um eine Vorbereitung auf ein Einstellungsgespräch geht (im präsentierten Text der erste Abschnitt) und somit zwei Abschnitte dahingehend irrelevant sind und ignoriert werden können. Weiterhin geht es um die Beschaffung von Informationen über den Arbeitgeber, die sich auf die ausgeschriebene Stelle beziehen. Der Text enthält auch in dem relevanten Abschnitt eine relativ große Anzahl konkurrierender Informationen (bspw. der Verweis auf die Internetseite des Unternehmens) und irrelevante Informationen (bspw. Informationen zum Ablauf des Gesprächs und betreffend der Nachbereitung). Diese müssen zunächst vom Probanden gelesen und dann entsprechend mental klassifiziert werden. Sodann gibt es drei verbleibende relevante Informationen:

1. Informationen über den Arbeitgeber bezüglich der Produkte, der Methoden oder Verfahren des Unternehmens = auf der Internetseite des Unternehmens oder bei der Handelskammer
2. Informationen über die Arbeitsstelle bezüglich der Ersetzung von jemandem oder Besetzung einer neu eingerichteten Position, ebenfalls über den Einsatz in welchen Abteilungen oder Geschäften = bei der Bundesagentur für Arbeit (BA)
3. Erwartungen des Unternehmen an potentielle Stelleninhaber = in der Stellenanzeige

Diese Informationen sind im Text strukturell nicht einfach aufbereitet, beispielsweise finden sich die Informationen, die bei der BA erhalten werden können, vor der Nennung der Quelle.

Überdies muss die konkrete Frage analysiert werden – es geht um Informationen, die von der BA zur Verfügung gestellt werden. Hierbei muss zunächst die im Text etablierte Abkürzung dechiffriert bzw. lokalisiert und identifiziert sowie die korrekte Angabe aus dem Text entnommen werden. Dazu kommt, dass die Distraktoren keine exakte Wiederholung des Textinhalts beinhalten, sondern eine leichte Selektion und Paraphrasierung darstellen.

Das folgende Item (Abbildung 19 und Abbildung 20) gehört zur Anforderungsklasse der *Generieren* Items und wurde als hybrider Stimulus klassifiziert. Diesbezügliche Erläuterungen folgen nach der allgemeinen Einführung des Items. Es erfordert die Transformation eines deskriptiven Textes in eine depiktionale Repräsentation (Skizze).

Unfallbericht L0649131

Bitte lesen Sie den Unfallbericht und ordnen Sie anschließend dem Bericht die richtige Skizze zum Unfallhergang zu.

Unfallbericht
Das Fahrrad fuhr auf der Vorfahrtstraße. Ein von rechts kommender PKW bog ohne Beachtung der Radfahlerin links in die Vorfahrtstraße ein. Die Radfahlerin musste abrupt bremsen und stürzte auf die Fahrbahn. Durch die Vollbremsung des auf der Vorfahrtstraße nachfolgenden Fahrzeugs kam es zu dem Auffahrunfall.

Skizze 1

Skizze 2

★ = Fahrradfahlerin

Abbildung 19 – Item 64 Seite 1

den Leser eine Vereinfachung darstellen sollen, als eine Widerspiegelung der tatsächlichen Gegebenheiten. Das Symbol für den Radfahrer ist völlig arbiträr und wird daher in der Legende erläutert – so simpel es scheinen mag, auch muss der Betrachter einen Zusammenhang zwischen der Legende und den Skizzen herstellen. Eine Handlung, die über das Lesen des Textes hinausreicht, wurde durch den ersten Satz induziert: „Bitte lesen Sie den Unfallbericht, um...“.

Beispielhaft deutlich wird an diesem Item auch die Schwierigkeit zwischen Itemteilen wie Hinführung, Stimulus, Frage und Antwort zu trennen. Zunächst erfolgt bei den Items eine kurze (1-2 Zeilen) Hinführung in die Situation. Da dies bei den meisten Items der Fall und aus Gründen der Logik auch erforderlich ist – denn nur so kann in diesem Format die Frage oder der Arbeitsauftrag übermittelt werden – enthalten nahezu alle Items einen deskriptionalen Anteil. Der eigentliche Stimulus (das „Arbeitsmaterial“) selbst kann depiktional, hybrid oder deskriptional sein, wird dann aber zumeist wieder geschlossen durch eine deskriptional formulierte Frage oder Aufforderung. Auch für das Antwortformat gilt, dass es deskriptional oder depiktional sein kann. Da bei der Klassifizierung der Stimulus als entscheidendes Merkmal ausschlaggebend berücksichtigt werden sollte, stellt dies eine gewisse Problematik dar, da die Trennung von Stimulus, Hinführung und Format der Antworten nicht immer trennscharf vorzunehmen ist, da jedes Items in einem gewissen Maß deskriptionale Elemente enthält. Bei der Konstruktion und Weiterentwicklung diesbezüglicher Instrumente sollte dies berücksichtigt werden. In diesem Fall besteht der Stimulus sowohl aus dem deskriptionalen Bericht in Form eines kurzen Fließtextes und verschiedenen Skizzen, die mehr oder weniger korrekt den Inhalt des Textes visualisieren. Davon abzugrenzen sind die gegebenen Antwortmöglichkeiten wieder in Form fragmentarischer deskriptionaler Repräsentationen (Teilsätze bzw. Stichworte). Würde man direkt die richtige Skizze ankreuzen lassen, könnte man das ansonsten unveränderte Item auch als rein deskriptional klassifizieren, da es so dann im Stimulus keine depiktionalen Anteile mehr gibt.

Bei diesem Item kommt erschwerend hinzu, dass es sich nicht auf einer Seite abbilden ließ, so dass die Probanden gezwungen sind zu blättern⁸⁹, was weitere Einflüsse durch Erinnerungsleistungen leider nicht völlig ausschließt.

Somit wird erkennbar, dass trotz großer Sorgfalt, die Items (alle) zwei Probleme aufweisen, die unter den gegebenen Rahmenbedingungen nicht zu lösen sind:

1. Es ist keine echte „Lesen, um zu Handeln“-Situation. Es bleibt durch den Testkontext eine „Lesen, um eine Leseaufgabe zu erfüllen“-Situation. Dies muss unbedingt durch die Weiterentwicklung von Testszenarien und die Integration von konkreten Handlungen in ebendiese Szenarien verbessert werden.
2. Die Fixierung auf das Single-Choice-Format ist insofern problematisch, da bspw. bei dem oben gezeigten Item de facto nur eine Abwägung von möglichen Visualisierungen

⁸⁹ Vor- und Zurückblättern innerhalb eines Items ist möglich – zwischen verschiedenen nicht

und die Entscheidung für die korrekte Antwort getroffen werden muss. Zwar muss auch zur Bewertung der Skizzen ein Moduswechsel erfolgen, wesentlich schwieriger und authentischer wäre es bei der Verwendung eines offenen Antwortformats, in diesem Fall, wenn die Skizze selbst angefertigt werden müsste. Die dann evtl. verworfenen Entwürfe und Korrekturen in der Skizze könnten zudem wichtige Hinweise auf die kognitiv stattfindenden Transformationsprozesse liefern.

Somit sind alle aus diesen Items abgeleiteten Folgerungen unter einer kritischen Perspektive zu betrachten und sollten nochmals unter adäquateren Testbedingungen überprüft werden.

Dieses Item gehört wie geschrieben in die Kategorie des *Generierens*. Ebenfalls zu dieser Klasse gehört das folgende Item, welches jedoch aus einer depiktionalen Repräsentation eine deskriptionale Entsprechung fordert:

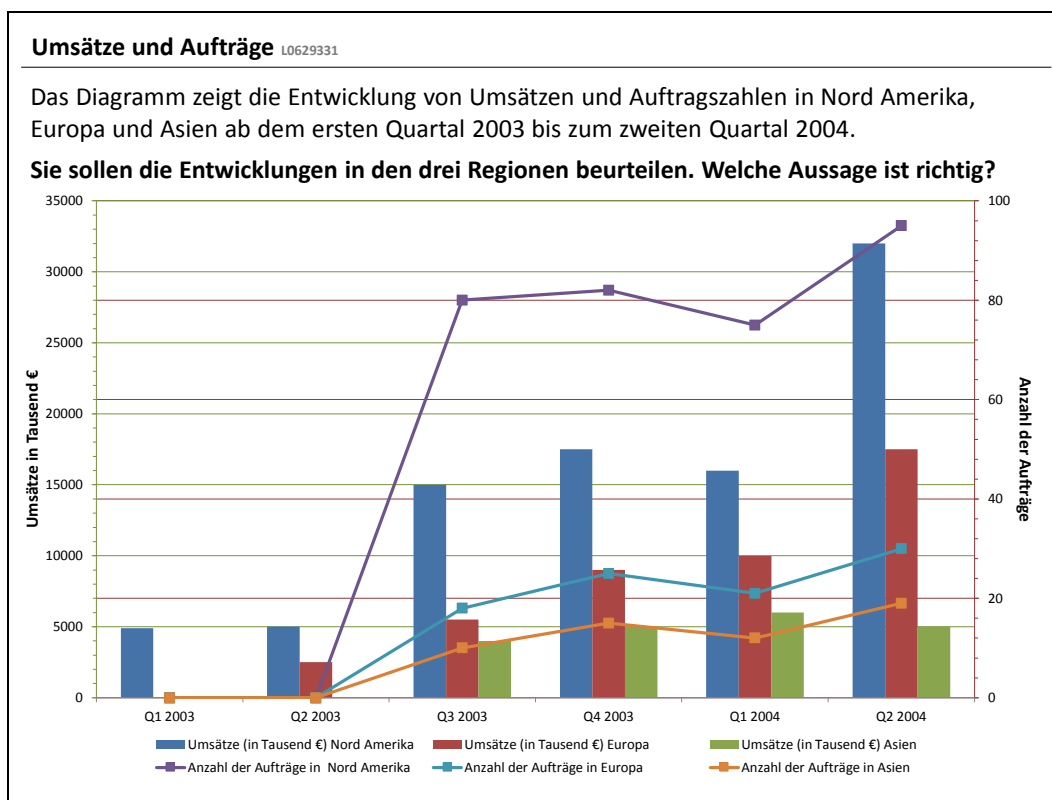


Abbildung 21 – Item 62 Seite 1

Welche Aussage zu den Entwicklungen ist richtig?

Bitte kreuzen Sie an.

- ☐ In allen Regionen kam es im ersten Quartal 2004 zu Umsatzrückgängen.
- ☐ In allen Regionen ist das Verhältnis von Auftragszahlen zu den Umsätzen ähnlich.
- ☐ In allen Regionen ist die Entwicklung der Umsätze ähnlich verlaufen.
- ☐ In allen Regionen ist die Entwicklung der Auftragszahlen ähnlich verlaufen.

Abbildung 22 – Item 62 Seite 2

In Item 62 wird die Handlungsintention durch einen Arbeitsauftrag induziert. Der Proband soll an Hand des vorliegenden Diagramms eine zu beobachtende Entwicklung ablesen, die Daten interpretieren und anschließend kurz und präzise formulieren. Dazu muss zunächst erkannt werden, dass das Diagramm zwei verschiedene Sachverhalte in ein und demselben Koordinatensystem abbildet – einerseits die Umsätze und andererseits die Auftragszahlen – mit jeweils unterschiedlichen Skalenniveaus: Umsätze in 1000€ auf der linken y-Achse und absolute Zahlen für die Aufträge auf der rechten Achse. Somit müssen die Linien und Balken korrekt den jeweiligen Aussagen zugeordnet werden. Da dieses Diagramm viele Aussagen bezüglich der Entwicklung in den einzelnen Regionen zulässt, von denen einige jedoch keinen Hinweis auf die Verarbeitungstiefe geben, wurden vier mögliche Aussagen formuliert, welche auf ihre Korrektheit geprüft werden sollen. Nur eine ist zutreffend, diese soll gewählt werden.

Item 62 ist hier nicht als hybrid zu bewerten, da zwar – wie bereits problematisiert – auch ein deskriptionaler Hinführungssatz vorliegt, jedoch der Stimulus eindeutig depiktional ist. Hieran wird auch nochmals der Unterschied zwischen dem Stimulus und dem Antwortformat deutlich. Parallel zum Item Nr. 64 liegt hier die Formulierung der Lösungen (Item 64: in Form von Bildern) bereits vor, allerdings knüpft daran keine Antwort mehr an, sondern die Auswahl einer Alternative stellt die Antwort da. Dass dies insgesamt nicht zufriedenstellend trennscharf vor allem bezüglich der dahinter ablaufenden kognitiven Prozesse ist, wurde bereits angemerkt.

Zur Beantwortung müssen die verschiedenen Aussagen zunächst verstanden werden und ihre mögliche visuelle Umsetzung antizipiert werden. Anschließend muss der Proband das Diagramm daraufhin untersuchen, ob eine der möglichen Umsetzungen zutrifft und ob Detailinformationen dem nicht widersprechen. Eine exakte Betrachtung fast aller Faktoren ist dazu notwendig. Dabei kommt es auf die Betrachtung globaler Zusammenhänge an, eine rein singuläre Fokussierung auf Detailvorgänge führt nicht zur richtigen Lösung.

Dies waren zwei vermutlich auf Grund ihrer Komplexität und der Vielzahl der durchzuführenden Schritte sehr anspruchsvolle Items. Somit wurde ein breites Spektrum an Anforderungen in den unterschiedlichen Repräsentationsformaten und kognitiven Anforderungsbereichen abgedeckt.

4.2. Analyse des Testinstruments

4.2.1. Kalibrierung der Testitems – empirische Schwierigkeitsbestimmung an Hand der Lösungswahrscheinlichkeit

Erkenntnisperspektive und Untersuchungsdesign

Wie bereits aufgeführt bewegt sich der empirische Teil der Untersuchung in einem Spannungsfeld zwischen der Entwicklung eines Instruments, welches verschiedenen aus den Einsatzfeldern und der Praxis gestellten Anforderungen gerecht werden und andererseits auch den aus der Theorie herausgearbeiteten Überlegungen und Konzeptionen Rechnung tragen soll.

Die zentrale *Erkenntnisperspektive*, die im Folgenden untersucht wird, sind die Faktoren „kognitiver Anforderungsbereich“ und „Repräsentationsformat“. Daraus folgend zwei Forschungsfragen:

1. Bilden sich die kognitiven Anforderungsbereiche in den ermittelten empirischen Schwierigkeiten der CBA-basierten Testung ab?
2. Gibt es Unterschiede in den empirischen Schwierigkeiten abhängig vom Repräsentationsformat (deskriptional/hybrid/depiktional)?

Dazu folgt eine ausführliche Betrachtung der durchgeführten *Kalibrierung* der Testitems – zur besseren Verständlichkeit mit einem kurzen Überblick über die Item-Response-Theory.

Ermittelt werden sollen statistisch signifikante Unterschiede mittels einer varianzanalytischen Betrachtung der mittleren empirischen Schwierigkeiten der Gruppen innerhalb jeder Kategorie – Kategorie meint hierbei die kognitiven Anforderungsbereiche bzw. das Repräsentationsformat.

Da die Testkonstruktion auf Basis der Probabilistische Testtheorie erfolgte, seien an dieser Stelle vorab kurz die wesentlichen Grundzüge der „Item-Response-Theory“ (IRT) oder „Probabilistische Testtheorie“ skizziert. Die Analysen erfolgten an Hand eines eindimensionalen Rasch-Modells, daher wird auch dieses kurz erläutert.

Item-Response-Theory/Probabilistische Testtheorie und Rasch-Skalierung

Grundsätzlich gibt es zwei messtheoretische Ansätze der Diagnostik – einerseits die „Klassische Testtheorie“ (KTT) und die „Item-Response-Theory“ (IRT) bzw. „Probabilistische Testtheorie“.

Die Klassische Testtheorie (KTT) wird auch als Messfehler-Theorie bezeichnet, da dies die zugrundeliegende Kernüberlegung der KTT ausdrückt. Diese besteht darin, dass sich der Messwert einer Person in einem Testitem immer aus zwei Komponenten zusammensetzt: der tatsächlichen Ausprägung des durch das Item erfassten Merkmals und einem zufälligen Messfehler (vgl. Moosbrugger 2012b, S. 104). Bereits hier wird das erste Problem der KTT deutlich, denn diese Annahme ist empirisch nicht überprüfbar, da beide Größen nicht direkt beobachtbar sind (vgl. Moosbrugger 2012b, S. 115). Zentraler Punkt der KTT ist, dass zur Bestimmung des wahren Werts bzw. um den Messfehler möglichst auszugleichen mehrmals gemessen wird und ein Mittelwert gebildet wird. Hierdurch müssten sich die zufälligen Fehler gegenseitig neutralisieren. Bei Persönlichkeits- und Leistungstests sind Wiederholungen des Messvorgangs mit demselben Messinstrument jedoch u.a. auf Grund von Erinnerungseinflüssen problematisch. Somit ist der wiederholte Einsatz desselben Instruments nicht zweckmäßig. Dieses Problem kann durch mehrere Messungen bei jedem Probanden mit verschiedenen Items gelöst werden. Diese müssen indes so konstruiert werden, dass alle dasselbe Konstrukt erfassen. Durch die Zusammenführung der Messungen mit den verschiedenen Items in einen Wert erfolgt theoretisch auch an dieser Stelle eine Neutralisation der Zufallsfehler (vgl. Moosbrugger 2012b, S. 106).

Insbesondere bleibt bei der KTT fraglich, ob die Annahme, dass die Reaktionen auf verschiedene Items zu einem Testwert zusammengefasst werden dürfen, gerechtfertigt ist. Denn dies setzt voraus, dass alle Items dasselbe Merkmal messen, was jedoch im Rahmen der KTT empirisch nicht überprüft werden kann (vgl. Moosbrugger 2012a, S. 228).

Im Gegensatz zur KTT werden bei der IRT die Antwortmuster jedes Items begutachtet. Bei der KTT ergibt sich wie oben ausgeführt die Fähigkeitsausprägung eines Probanden durch das Gesamtergebnis – dagegen können in der IRT Beziehungen direkt zwischen einzelnen Testitems und einer latenten Variablen (bspw. der Lesekompetenz) ermittelt werden (vgl. Steyer/Eid 1993, S. 215). Dies ist nach Bühner auch der Hauptunterschied zwischen KTT und IRT: in der KTT wird angenommen, dass der Summenwert einer Person in einem Test einen Messwert darstellt. Im Gegensatz dazu wird im Rahmen der IRT berücksichtigt, ob es sich bei der Summation der Items überhaupt um eine gültige Verrechnungsvorschrift handelt (vgl. Bühner 2006, S. 300). Dies umgeht das oben beschriebene Problem der KTT.

Sodann können mit der IRT verschiedene Fragestellungen geklärt werden. Es kann einerseits getestet werden, ob das zu messende Merkmal durch alle Testitems kontrolliert wird und andererseits in welchem Ausmaß dabei ein Messfehler berücksichtigt werden muss. Die Testgütekriterien wie Reliabilität, Validität und Objektivität können für Gewöhnlich bei der Geltung eines probabilistischen Testmodelles für einen Datensatz als erfüllt angenommen werden, da diese innerhalb der Modellgeltungstests mit überprüft werden (vgl. Bühner 2006, S. 342). Das zentrale Kriterium Eindimensionalität des Tests, was meint, dass die Itemantworten nur durch eine latente Variable bedingt sind (vgl. Bühner 2006, S. 300), kann direkt getestet werden. Sofern das probabilistische Testmodell bzgl. der ermittelten Daten als gültig angenommen wird, ist der Test auch tatsächlich eindimensional (vgl. Jordan 2011, S. 133).

Die beobachteten Antwortmuster, sogenannte Pattern, müssen gemäß der IRT einem bestimmten Testmodell folgen, wobei verschiedene theoretische Grundlagen solcher Testmodelle existieren⁹⁰. Eine der Grundannahmen der IRT liegt darin, dass mit ansteigender Personenfähigkeit eine korrekten Itemlösung probabilistisch (daher auch die Bezeichnung probabilistische Testtheorie) zunimmt. In diesem Fall wird von kontinuierlich ansteigenden Itemfunktionen gesprochen. Diese Annahme kann durchaus auch für die KTT getroffen werden, jedoch kann mit Hilfe der KTT keine Aussage über die Lösungswahrscheinlichkeit eines Items für eine bestimmte Person vorhergesagt werden – ganz im Gegensatz zur IRT: Mittels einer Modellgleichung kann über die IRT die Lösungswahrscheinlichkeit jedes Items für eine bestimmte Person ermittelt werden, da in der IRT die Lösungswahrscheinlichkeit eines Items von zwei Faktoren abhängt: einerseits von der Fähigkeits- oder Eigenschaftsausprägung einer Person, andererseits von der Schwierigkeit des jeweiligen Items – sind beide bekannt, können sie ohne weiteres in die Modellgleichung eingesetzt werden (vgl. Bühner 2006, S. 33; Jordan 2011, S. 133).

Es gibt unterschiedliche Verfahren, um die Relation zwischen Fähigkeitsausprägung der Testpersonen und Lösungswahrscheinlichkeit mathematisch zu modellieren. Nach Jordan (2011) ist das innerhalb von Testentwicklungen am häufigsten eingesetzte probabilistische Testmodell das dichotome Rasch-Modell – da dieses später auch zur Analyse des vorliegenden Datensatzes eingesetzt wird und wichtige Informationen über die Items liefert, wird es kurz näher betrachtet. Die Annahme eines sprunghaften Anstiegs würde bedeuten, dass insbesondere Personen, die vom Fähigkeitsindex am Rand einer Sprungstelle bzw. auf beiden Seiten derselben liegen, sich von der Fähigkeitsausprägung kaum unterscheiden, jedoch innerhalb der Itemfunktion einer Sprung-Skala eine deutliche Unterscheidung zwischen diesen Personen gemacht würde. So löst ein Proband mit einer bestimmten Personenfähigkeit ein Item mit einer Wahrscheinlichkeit von $p=1$, ein Proband mit einer nur minimal nach unten abweichenden Fähigkeit kann das Item laut Theorie aber nicht lösen, obwohl beide Probanden in etwa dieselbe Fähigkeit haben. Dies scheint nur bei wenigen Leistungsanforderungen die Realität angemessen abzubilden. Einen weitaus besseren Erklärungsansatz und Modellabbildung bietet eine monotone, stetige

⁹⁰ vgl. dazu ausführlich Rost 2004, S. 89–300.

Itemfunktion – wie das Rasch-Modell, bei dem nur im Mittelbereich der Skala ein linearer Zusammenhang zwischen der Fähigkeit der Person und der Lösungswahrscheinlichkeit der Person bzgl. eines bestimmten Items angenommen wird (vgl. Jordan 2011, S. 137).

Wie die meisten Testmodelle geht das Rasch-Modell auf Grund verschiedener Argumente somit statt von einem sprunghaften Anstieg der Lösungswahrscheinlichkeit von einer monoton ansteigenden, stetigen Itemfunktion aus – also einer Itemfunktion ohne Sprungstellen aus.

Die Grundidee des Rasch-Modells ist folgende: für jedes Item wird in Abhängigkeit des zu schätzenden Itemparameters eine Itemfunktion bestimmt. Sodann kann für jede Person abhängig vom zu schätzenden Personenparameter durch Einsetzen in die Modellgleichung (oder durch Ablesen des Y-Wertes der Itemfunktion) die Wahrscheinlichkeit bestimmt werden, mit der das jeweilige Item gelöst wird. An Hand der Modellgleichung⁹¹ ist zu sehen, dass sich Itemfunktionen verschiedener Items innerhalb des Rasch-Modells nur durch Parallel-Verschiebungen bzgl. der X-Achse unterscheiden (vgl. Rost 2004, S. 120). Dabei liegen die Personenparameter und die Itemparameter auf derselben Skala bzw. in derselben Einheit vor und besitzen Differenzskalenniveau – wobei der jeweilige Wertebereich theoretisch von - bis + reichen kann, jedoch in der Regel zwischen -3 und +3 liegt (vgl. Rost 2004, S. 121; Bühner 2006, S. 314; Jordan 2011, S. 138f).

Beim Rasch-Modell wird in die theoretischen Überlegungen einbezogen, dass die Itemfunktion nur im Mittelbereich einen annähernd linearen Verlauf hat – für die x-Werte Richtung positiv/negativ unendlich nähert sich die Funktion asymptotisch der 1 bzw. der 0. Der S-kurvige Funktionsverlauf erscheint insbesondere bei Leistungstests psychologisch plausibel, da durch ihn die Annahme beschrieben wird, dass die Lösungswahrscheinlichkeit im Mittelbereich am stärksten mit zunehmender Fähigkeit steigt. Wenn ein Item dagegen im Vergleich mit den Fähigkeiten zu leicht oder zu schwer ist, verändert eine geringfügige Fähigkeitssteigerung die Lösungswahrscheinlichkeit nur minimal (vgl. Rost 2004, S. 115).

Das heißt für die Testerstellung, dass zunächst die Itemkennwerte ermittelt werden müssen. Dies erfolgt im Rahmen einer Kalibrierungsstudie, die im Folgenden erläutert wird.

Durchführung und Ergebnisse der Kalibrierung

Die Kalibrierung erfolgte wie auch die Konstruktion der Items im Rahmen der ASCOT-Initiative (vgl. ausführlich Kap. 4.1.2), daher wurden auch zwei andere Domänen (Mathe und Naturwissenschaften) mit erhoben. Wie bereits ausgeführt erfolgte die Erhebung computerisiert. Dabei wurde ein sogenanntes Balanced-Block-Design verwendet, um einerseits Positionseffekte auf Itemebene zu verhindern und zu ermöglichen, dass alle Items von ungefähr der gleichen Anzahl von Probanden beantwortet werden. Balanced-Block-Design bedeutet hier, dass nicht allen Pro-

⁹¹ $p(x_{vi}) = \frac{\exp(x_{vi} \times (\theta_v - \sigma_i))}{1 + \exp(\theta_v - \sigma_i)}$

x_{vi} = Antwort der Person v bei Item i ; 1=Item wurde korrekt gelöst; 0=Item wurde nicht korrekt gelöst;

θ_v = Personenfähigkeit der Person v ; σ_i = Itemschwierigkeit des Items i

banden alle Fragen vorgelegt werden, stattdessen werden Blöcke gebildet – in diesem Fall entsprechend der drei Domänen (Lesen, Mathe, Naturwissenschaften), deren Position rotieren. Jeder der drei Blöcke wird jedem Probanden vorgelegt und innerhalb dieser Blöcke werden jedem Probanden eine bestimmte Anzahl Fragen (Lesen: 9, Mathe und Naturwissenschaften: 12) vorgelegt. Somit erhielt einerseits jeder Proband die ungefähr gleiche Anzahl an Fragen und jede Frage wurde von ungefähr der gleichen Anzahl Probanden beantwortet, jedoch ist diese Zahl logischerweise wesentlich kleiner als die Gesamtstichprobe. Tabellarisch darstellen lässt sich das Testheftdesign wie folgt:

Anordnung 1	Anordnung 2	Anordnung 3	Anordnung 4	Anordnung 5	Anordnung 6
Lesen	Lesen	Nawi	Nawi	Mathe	Mathe
Mathe	Nawi	Lesen	Mathe	Nawi	Lesen
Nawi	Mathe	Mathe	Lesen	Lesen	Nawi

Abbildung 23 – Domänenfolge der Testhefte (Bernhard 2012)

Innerhalb einer Anordnung gibt es auf Grund der Itemanzahl der drei Domänen 133 verschiedene Kombinationen von Items, d.h. insgesamt $(133 \cdot 6)$ 798 verschiedene Testhefte.

Die Kalibrierungsstudie wurde nach Überprüfung der technischen Konstellationen und der Durchführbarkeit an den jeweiligen Schulen von September bis November 2012 9 Wochen lang mit durchschnittlich 2,5 Testtagen pro Woche ausgeführt.

Insgesamt wurde die Kalibrierung in den Bundesländern Hessen, Thüringen und Niedersachsen durchgeführt. Die Gesamtstichprobe umfasste 1.632 Schüler⁹². Davon wurden in Hessen an 9 Schulen 475 Schüler getestet.

46,3% gaben an weiblich zu sein, 52,6% männlich und 1,2% machten keine Angabe.

Das Alter der Probanden lag im Mittel bei 20 (Standardabweichung: 3). Das Minimum laut Angabe war 16 und das Maximum 43, keine Angabe machten 3,7%, damit lag der Anteil der 18-22 Jährigen – also die primäre Zielgruppe: Jugendliche zum Ende ihrer Ausbildung – bei 69%.

Die hauptsächlich gesprochene Sprache war zu 86,9% Deutsch.

Der höchste Schulabschluss der Probanden war in 28,5% der Fälle Abitur, in 62,1% Mittlere Reife und in 7,2% Hauptschulabschluss (keine Angabe machten 1,4%).

Die getesteten Berufsbereiche setzten sich zu 22,9% aus medizinisch-pflegenden Berufen (MFA, Altenpfleger, ZMFA,...), zu 35,8% aus gewerblich-technischen Berufen (diverse Elektronikberufe, Metall- und Technikberufe,...) zu 33,8% aus kaufmännisch-verwaltenden Berufen (diverse Kaufleute, Steuerfachangestellte,...) und zu 5,5% aus sonstigen Berufen (Gärtner, Förster, ...) zusammen – 2,5% der Probanden machten keine Angaben zu ihrem Beruf.

⁹² Damit wurde jede mögliche Itemkombination (Testheft) von mind. 2 Schülern bearbeitet.

Nach der Erhebung und Bereinigung der Daten, wobei die Items auf geschlechtsbezogene Diff-Effekten⁹³ untersucht und nach einer inhaltlichen Prüfung 8 Items (ID: 22, 26, 30, 31, 44, 53, 65 und 71) ausgeschlossen wurden⁹⁴, erfolgte eine Analyse sowohl der Testfragen als auch des Tests. Dabei wurden die Items eindimensional gemäß dem Raschmodell (einfaches logistisches Modell) mittels Gauss-Hermite Quadratur ohne Gewichtungen skaliert, d.h. Itemschwierigkeit und Personenfähigkeit werden auf einer Skala (graphisch bspw. in Form einer „Latent Distribution Map“ (vgl. Wu et al. 2007, S. 25)) dargestellt. Dabei sind die Items entsprechend der jeweilig ermittelten Schwierigkeit (bzw. Lösungswahrscheinlichkeit von $p=.5$ bei einem bestimmten niedrigen oder hohen Personenparameter) geordnet. Im unteren Bereich der Skala sind Items mit geringer Schwierigkeit bzw. Personen mit geringer Fähigkeit zu finden und im oberen Bereich schwerere Items/Personen mit höheren Personenfähigkeiten (vgl. bspw. Abbildung 24 für diesen Test). In dieser Abbildung werden Personen durch ein ‚x‘ symbolisiert, jedoch erfolgt die Darstellung nicht 1:1. Daher wird für gewöhnlich ebenfalls angegeben wie viele Fälle einem ‚x‘ entsprechen (hier: $x = 2.3$ Fälle). Die auf der rechten Seite abgetragenen Zahlen stehen für die Items, die je nach ihrer Schwierigkeit entlang der Skala positioniert sind⁹⁵. Die Personenfähigkeiten und die Itemschwierigkeiten werden wie bereits geschrieben auf einer gemeinsamen Skala abgebildet – einer sogenannten Logit-Skala. Hierzu ist in aller Kürze auszuführen, dass sich die mathematische Einheit der Rasch Messung, die „log-odds unit“ bzw. „logit“ aus dem Abstand von Personenfähigkeit θ und Itemschwierigkeit δ durch den Logarithmus des Quotienten der Wahrscheinlichkeit, das Item richtig zu lösen, und der Wahrscheinlichkeit, falsch zu lösen, ergibt – sie kann theoretisch Werte zwischen $-\infty$ und ∞ annehmen (vgl. Wu et al. 2007, S. 134f). In diesem Fall liegt der Großteil der Items zwischen -2 und +2 (vgl. Abbildung 24).

Wie ebenfalls an Hand der Abbildung 24 zu erkennen ist, sind die Items insgesamt über die Skala verteilt.⁹⁶ Es ist erkennbar, dass sowohl besonders leichte (unten) als auch besonders schwere Items (oben) gibt. Dies zeigt jedoch auch, dass das durch die Items abgebildete Spektrum sehr breit gefächert ist. Tendenziell als problematisch zu bewerten ist die größere Lücke zwischen 2 Itemblöcken im unteren Bereich (markiert durch den blauen Chevron).

⁹³ Die Analysen wurden von den Projektpartnern Andreas Frey und Raphael Bernhardt durchgeführt. Berechnet wurde das Modell $\text{item} + \text{gender} + \text{item} * \text{gender}$ (weitere Ausführungen siehe: Bernhardt und Frey 2013, Spoden et al. 2014, 2015).

⁹⁴ Bei allen anderen Items gab es bezüglich dieser möglichen Problematik keine Auffälligkeiten.

⁹⁵ Die Nummern bezeichnen die in der jeweiligen Berechnung verwendeten Itemnummerierungen und entsprechen nicht den bei der Konstruktion ursprünglich definierten Item-IDs.

⁹⁶ Zur Erläuterung: bei einer latent distribution Map werden die Werte der Items und die der Probanden auf derselben Logit-Skala abgetragen. Die Nummern entsprechen den Items, die x entsprechen den Fähigkeitsparametern der Probanden. Diese können also über, unter und zwischen den Werten der Items liegen. Dies bedeutet, dass nicht allen Fähigkeiten exakt Items entsprechen. Besonders offenkundig wird dies an der mit einem blauen Chevron markierten Stelle. In diesem Bereich bewegen sich einige Personen, deren Wahrscheinlichkeit das Item darüber zu lösen bei unter 50% liegt, deren Wahrscheinlichkeit das darunterliegende Item zu lösen, jedoch bei über 50% liegt.

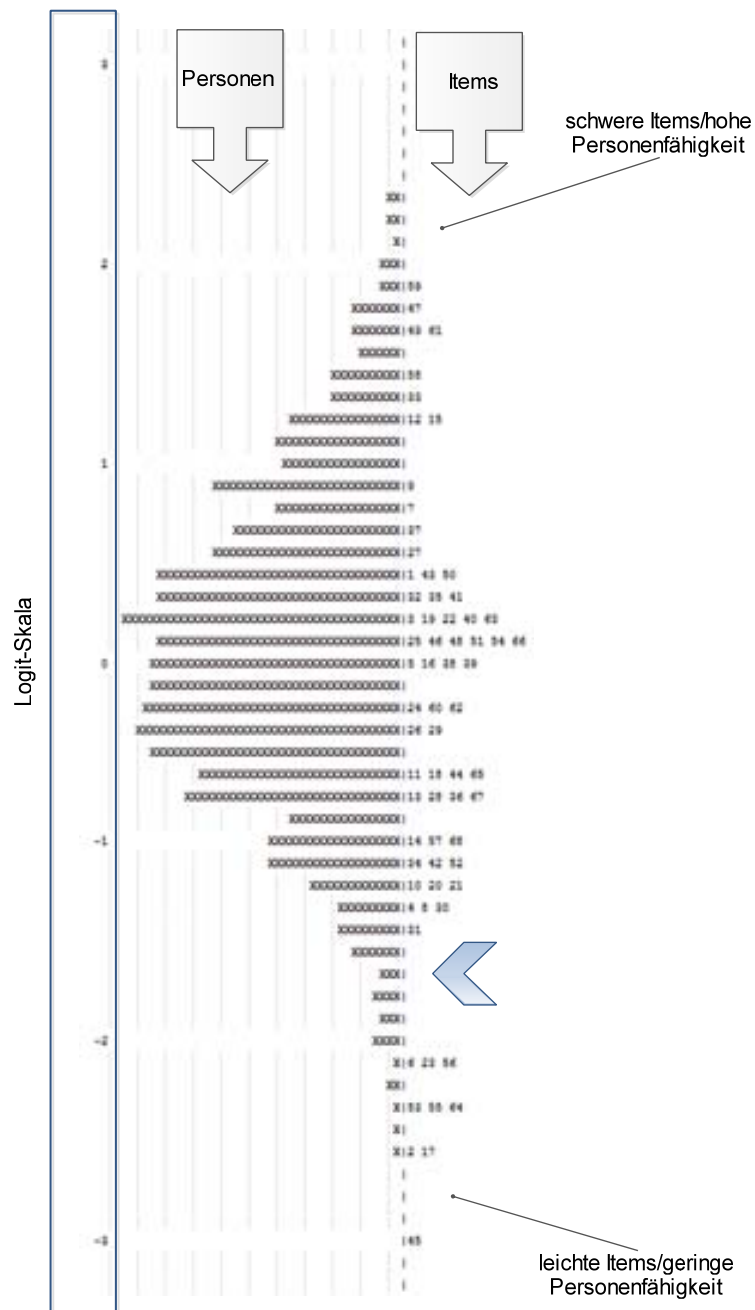


Abbildung 24 – Latent Distribution Map (CBA)

Auf Testebene zeigt sich ferner, dass die Berechnungen beendet wurden, da das Konvergenzkriterium erreicht wurde, d.h. die Lösungen sind vor 1000 Berechnungen konvergiert – genauer: nach 22 Berechnungen. Es gibt somit einen Grenzwert, was dahingehend interpretiert werden kann, dass eine optimale Lösung gefunden wurde und kein Abbruch auf Grund von definierten Kriterien (wie bspw. Anzahl der Berechnungsversuche) notwendig war.

Die sogenannte Test Characteristic Curve (TCC) entspricht der Zusammenfassung der itemcharakteristischen Funktionen (ICC – siehe unten) aller Items im Test. Wie zu sehen ist, entspricht diese zufriedenstellend einer leicht abgeflachten S-Form die keine auffällige Verschiebung nach

rechts oder links aufweist. Damit ist graphisch dargestellt, dass der Test sowohl im unteren als auch im oberen Kompetenzbereich Ergebnisse erzielt.

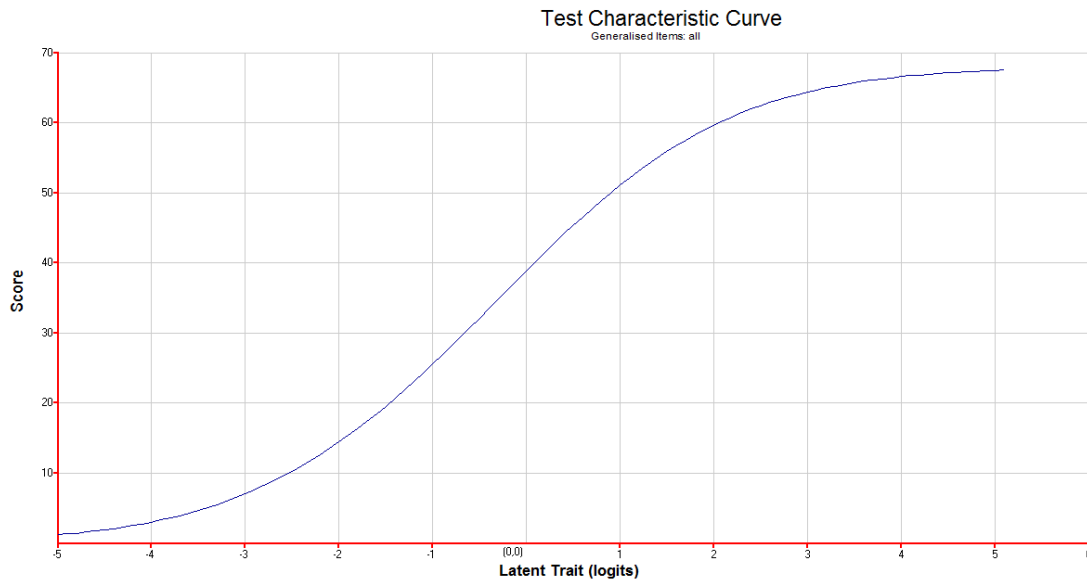


Abbildung 25 – Test Characteristic Curve (CBA)⁹⁷

Die Testinformationsfunktion (TIF) summiert gleichsam der TCC die einzelnen Iteminformationsfunktionen auf (IIF). Diese sieht für die betreffenden Items wie folgt aus:

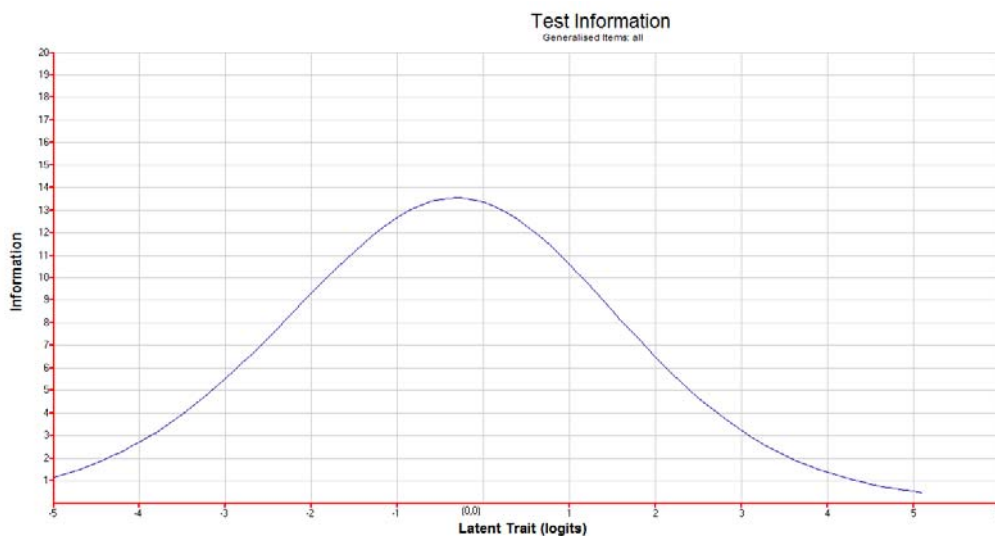


Abbildung 26 – Test Information Function (CBA)

Hier ist erkennbar, dass besonders trennscharfe Ergebnisse im Bereich von -2 bis ca. +1.5 erzielt werden. An der Stelle an der die Testinformationsfunktion ihren höchsten Punkt besitzt, ist die Diskriminationsfähigkeit des Tests am höchsten. Abbildung 26 zeigt, dass die Testinformations-

⁹⁷ CBA = Computerbased Assessment

funktion ihr Maximum entsprechend der Normalverteilungsannahme in der Nähe des Nullpunktes hat und mit zunehmender Abweichung vom Mittelwert der Informationsgehalt sinkt bzw. die Varianz zunimmt.

Die detaillierten Ergebnisse auf Itemebene sind Tabelle 6 zu entnehmen:

Tabelle 6 – Itemkennwerte (CBA)

ID	Estimate	Error	WMNSQ	t	Cases	DC	Time
01	.397	.163	1.00	.1	191	.43	01:32
02	-2.529	.266	.98	.0	188	.36	00:57
03	.196	.163	1.00	.1	186	.46	01:30
04	-1.292	.185	.99	-.1	191	.47	01:18
05	-.077	.159	1.08	1.5	194	.36	02:17
06	-2.103	.231	1.03	.3	191	.36	01:41
07	.717	.172	.99	-.1	181	.45	02:24
08	-1.253	.180	1.00	.0	198	.36	00:51
09	.864	.169	1.00	.0	194	.46	02:35
10	-1.161	.182	1.00	.0	188	.45	01:00
11	-.644	.163	.99	-.2	200	.47	02:56
12	1.147	.184	.99	-.1	181	.44	01:02
13	-.691	.172	.94	-1.0	180	.53	01:16
14	-.944	.175	1.05	.7	189	.34	01:51
15	1.151	.178	.90	-1.2	193	.56	01:40
16	-.042	.167	1.10	1.9	177	.35	02:03
17	-2.525	.261	.96	-.2	192	.42	01:05
18	-.660	.167	.93	-1.2	190	.57	01:54
19	.186	.160	1.09	1.7	196	.36	02:26
20	-1.177	.180	.98	-.2	191	.40	01:05
21	-1.184	.182	1.02	.3	188	.37	01:13
23	.138	.159	.95	-.9	198	.52	03:47
24	-2.069	.234	.92	-.5	181	.48	01:10
25	-.325	.171	.95	-.8	173	.53	01:59
27	-.404	.167	.94	-1.0	184	.54	02:09
28	.457	.161	.99	-.2	197	.47	01:12
29	-.685	.166	1.06	1.0	194	.38	02:33
32	-.432	.163	1.00	.1	192	.42	02:04
33	-1.267	.182	.93	-.7	196	.52	01:04
34	-1.409	.187	1.07	.7	192	.36	02:26
35	.252	.161	.98	-.4	193	.50	01:46
36	1.331	.186	1.02	.2	190	.33	01:13
37	-1.066	.179	.97	-.3	186	.53	01:23
38	.300	.170	1.01	.2	174	.44	01:56
39	-.732	.173	.90	-1.5	182	.65	01:41
40	.645	.167	1.01	.2	189	.44	01:59
41	-.072	.163	1.04	.8	186	.42	02:26
42	-.002	.171	1.08	1.6	167	.32	01:42

43	.129	.163	.90	-2.0	187	.59	01:14
45	-1.087	.182	.99	-.1	184	.43	00:58
46	.374	.160	1.02	.4	196	.43	01:31
47	-.582	.165	1.00	.1	191	.52	02:04
48	-2.972	.310	.98	.0	195	.29	01:09
49	.020	.161	.98	-.4	192	.51	01:40
50	1.742	.203	1.00	.0	188	.38	01:41
51	.011	.163	1.01	.1	185	.46	02:05
52	1.660	.202	1.00	.0	191	.37	01:21
54	.425	.161	1.01	.2	195	.48	01:23
55	.076	.160	.96	-.7	195	.50	01:29
56	-1.094	.178	.99	-.1	191	.47	01:10
57	-2.288	.247	.95	-.3	181	.44	02:05
58	.088	.160	1.01	.2	194	.44	02:17
59	-2.261	.252	.92	-.4	177	.53	00:50
60	-2.147	.231	1.00	.1	196	.33	00:50
61	-.899	.172	.97	-.4	191	.54	00:50
62	1.435	.191	.99	-.1	185	.41	01:40
63	1.823	.207	1.00	.1	200	.33	01:17
64	-.297	.166	1.05	.9	182	.39	01:38
66	-.276	.161	1.01	.2	195	.44	01:55
67	.184	.161	.99	-.3	194	.45	01:07
68	-2.259	.243	.88	-.7	184	.54	00:52
69	-.628	.164	.96	-.6	196	.54	01:21
70	.073	.160	1.01	.2	194	.48	01:43
72	-.748	.170	1.06	1.0	185	.38	01:23
73	-.960	.173	1.04	.5	192	.35	01:03

Die Tabelle ist folgendermaßen zu lesen:

Der *Estimate*wert gibt den Itemparameterwert (Punkt auf der Merkmalsskala mit 50% Lösungswahrscheinlichkeit), d.h. die „Schwierigkeit“ an und der Error den Standardfehler. Der *WMNSQ*⁹⁸ gibt die gewichteten Abweichungsquadrate bzw. den Itemfit an, wobei 1 als exakter Fit im Hinblick auf das gewählte Messmodell betrachtet wird und die absolute Werte nicht größer als 1.33 (vgl. Adams/Khoo 1996) bzw. nach konservativerer Betrachtung (bspw. PISA Technical Report) nicht größer als 1.20 sein sollten (vgl. Adams, S. 105). Der *t-Wert* wiederum stellt die Trennschärfe dar. Wenn t größer als 2 sein sollte, so liegt eine signifikante Abweichung in Richtung einer zu niedrigen Trennschärfe vor, negative Werte hingegen weisen auf eine zu hohe Trennschärfe hin und werden in der Regel nicht als problematisch angesehen. Unter *Cases* ist die Anzahl der Personen zu finden, die Antworten zu dem Item abgegeben haben (richtig/falsch – ohne Missings). Dies ist auf Grund des Balanced-Block-Designs hier von Interesse. Abschließend ist unter *DC* (= *Discrimination*) die klassische Trennschärfe, d.h. die Korrelation

⁹⁸ = weighted mean square; gewichtete Abweichungsquadrate

eines Items mit dem Gesamtergebnis des Tests, zu finden. Generell gilt, dass die Definitionsmenge zwischen -1 und +1 liegen kann. Bei einer hohen positiven Trennschärfe des Items wird davon ausgegangen, dass es etwas Analoges erfasst wie der Gesamttest. Eine Trennschärfe, welche nahe 0 liegt, weist dagegen darauf hin, dass das entsprechende Item mit dem restlichen Test wenig übereinstimmt. Eine negative Trennschärfe ist insofern genauer zu betrachten, da diese darauf hinweisen kann, dass ein Item entgegengesetzt als berechnet verstanden oder bei der Auswertung fehlerhaft gepolt wurde. Als Richtlinie bezüglich der Itemselektion an Hand der Trennschärfe gilt, dass die Items wenigstens eine Trennschärfe von .30 oder höher aufweisen sollten (Jude 23.02.2006).

Weiterhin ist unter *Time* die durchschnittliche Bearbeitungsdauer des jeweiligen Items in Minuten und Sekunden angegeben.

An Hand Tabelle 6 lässt sich Verschiedenes erkennen:

Beginnend mit der Analyse der *Cases* lässt sich festhalten, dass alle Items von mindestens 167 Personen beantwortet wurden, im Durchschnitt wurde ein Item ca. 189 Personen vorgelegt, und die Differenz zwischen dem meist beantworteten und dem am wenigsten beantworteten Item beträgt 33. Daher kann davon ausgegangen werden, dass einerseits das oben erläuterte Testheftdesign erfolgreich eingesetzt wurde und alle Items ungefähr gleich häufig vorgelegt wurden, andererseits kann dadurch von einer verhältnismäßigen Repräsentativität bezüglich der Stichprobe ausgegangen werden, d.h. es ist unwahrscheinlich, dass ein Item nur von männlichen Probanden im Beruf Mechatroniker beantwortet wurde.

Der *WMNSQ* weist bei keinem Item einen Wert größer als 1.20 auf, was zeigt, dass alle Items einen zufriedenstellenden Fit bezüglich des Modells aufweisen, 11 besitzen einen exakten Fit, der maximale Wert beträgt 1.1 (Item 16) und der minimale.88 (Item 68) – zu diesem ist zu sagen, dass er nicht kleiner als.75 sein sollte (Bond/Fox 2001), um noch als gut bewertet zu werden.

Der *t-Wert* für die Trennschärfe beträgt maximal 1.9 (Item 16) und ist damit sowohl gemäß liberaler Betrachtungen (<2) also auch nach konservativen (<1.96) über alle Items hinweg als gut zu beurteilen.

Bezüglich der *Discrimination* (DC) lässt sich festhalten, dass diese nur in einem Fall Werte unter.3 annimmt (Item 48), allerdings ist der Wert von.29 noch im akzeptablen Bereich und auf Grund des guten Fits und des guten t-Werts kein Ausschussgrund für das Item.

Die *Bearbeitungsdauer* der Items differiert von 50 Sekunden bis hin zu 3 Minuten und 47 Sekunden, was eine Differenz von fast 3 Minuten ist. Die querschnittliche Bearbeitungsdauer beträgt eine Minute und 38 Sekunden. Alle Items summiert haben eine Gesamtbearbeitungsdauer von einer Stunde und 45 Minuten. Durch das Testheftdesign kann davon ausgegangen werden, dass sowohl motivationale Effekte wie auch der Einfluss der Konzentration insgesamt gemittelt wird.

Zur den *Estimate-Werten* kann festgehalten werden, dass das Item mit der höchsten Lösungswahrscheinlichkeit ($\hat{=}$ das leichteste Item) einen Wert von -2.972 aufweist (Item 48) und das Item mit der geringsten Lösungswahrscheinlichkeit ($\hat{=}$ das schwierigste Item) einen Wert von 1.8 (Item 63) besitzt. Des Weiteren ist die querschnittlichen Schwierigkeit -.42. was leicht unter dem anzustrebenden 0 Punkt liegt.

An dieser Stelle lässt sich zusammenfassen, dass die Items insgesamt als zufriedenstellend bewertet werden können – sowohl was ihre verschiedenen Gütekriterien (Fitwerte, Trennschärfe,...) betrifft als auch bezüglich ihre Abdeckung eines breiten Spektrums an kognitiven Anforderungen, Schwierigkeiten und Repräsentationsformaten.

Da an dieser Stelle aus pragmatischen Gründen nicht alle 65 Items ausführlich besprochen werden können, werden die gleichen Items gewählt, die bereits in Kapitel 4.1.4 vorgestellt und in Kapitel 4.2.4 thematisiert wurden (Item 6, 49, 59, 60, 62 und 64).

Tabelle 7 – Itemkennwerte Item 6, 49, 59, 60, 62 und 64

ID	Estimate	Error	WMNSQ	t	Cases	DC	Time
6	-2.103	.231	1.03	.3	191	.36	01:41
49	.020	.161	.98	-.4	192	.51	01:40
59	-2.261	.252	.92	-.4	177	.53	00:50
60	-2.147	.231	1.00	.1	196	.33	00:50
62	1.435	.191	.99	-.1	185	.41	01:40
64	-.297	.166	1.05	.9	182	.39	01:38

Wie an Hand der *Estimates* zu sehen ist, sind die Items 6, 59 und 60 eher dem leichten, die Items 49 und 64 dem mittleren und Item 62 dem schwierigen Bereich zuzuordnen. Dies ergibt sich aus den Estimatewerten: geringe Werte bedeuten eine hohen Lösungswahrscheinlichkeit $\hat{=}$ leichte Items und hohe Estimatewerte bedeuten eine geringe Lösungswahrscheinlichkeit, d.h. schwierige Items. Der *WMNSQ* weist für alle sehr gute (6, 60 – exakt, 64) bis „überfittete“ (49, 59 und 62) Werte auf, desgleichen gilt für den *t-Wert*.

Da sich aus 6 Items keine allgemeinen Aussagen ableiten lassen, werden die Präsentationsformate sowie die kognitiven Anforderungsbereich unten nochmals detaillierter in Bezug zu den in der Kalibrierung ermittelten Schwierigkeiten gesetzt.

Bezüglich der Bearbeitungszeit fällt auf, dass diese kaum Hinweise auf die Schwierigkeit bietet. Sowohl das am längsten bearbeitete Item als auch das mit der kürzesten Bearbeitungszeit sind leichte Items.

Weiterhin können bezüglich der angesprochenen Items die item characteristic curves (ICC) betrachtet werden. Diese visualisieren den Zusammenhang zwischen dem latenten Merkmal (Personenfähigkeit) und der Wahrscheinlichkeit für eine bestimmte Lösung (Itemschwierigkeit) bezüglich des beobachteten Antwortverhaltens (vgl. Winther 2010, S. 128).

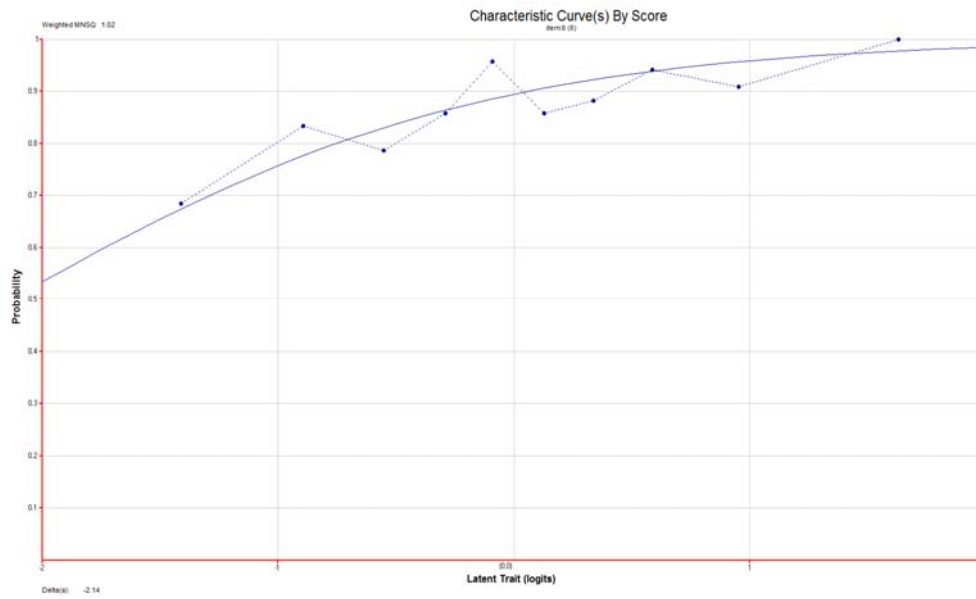


Abbildung 27 - ICC Item 6 (CBA)

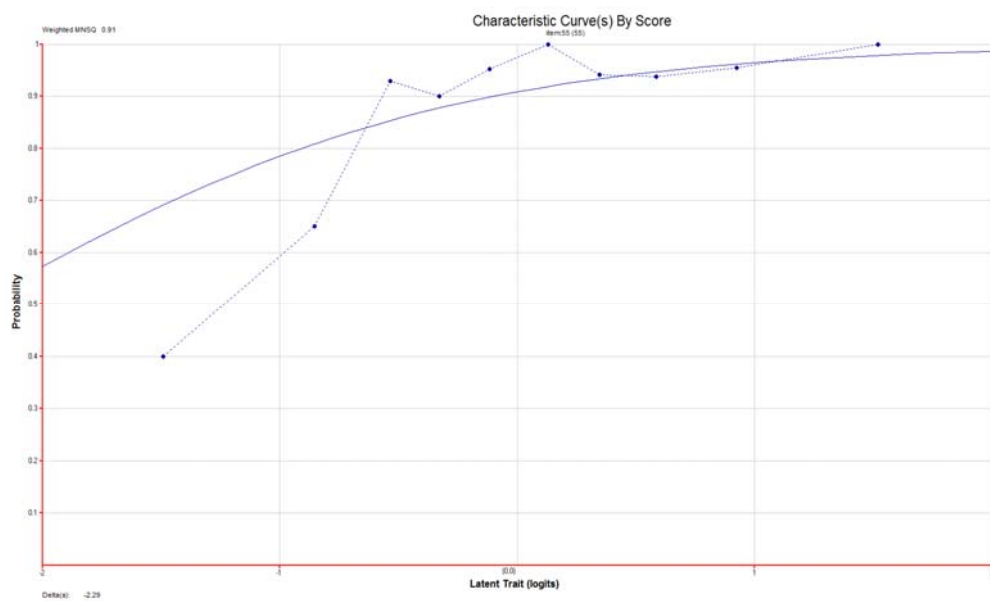


Abbildung 28 - ICC Item 59 (CBA)

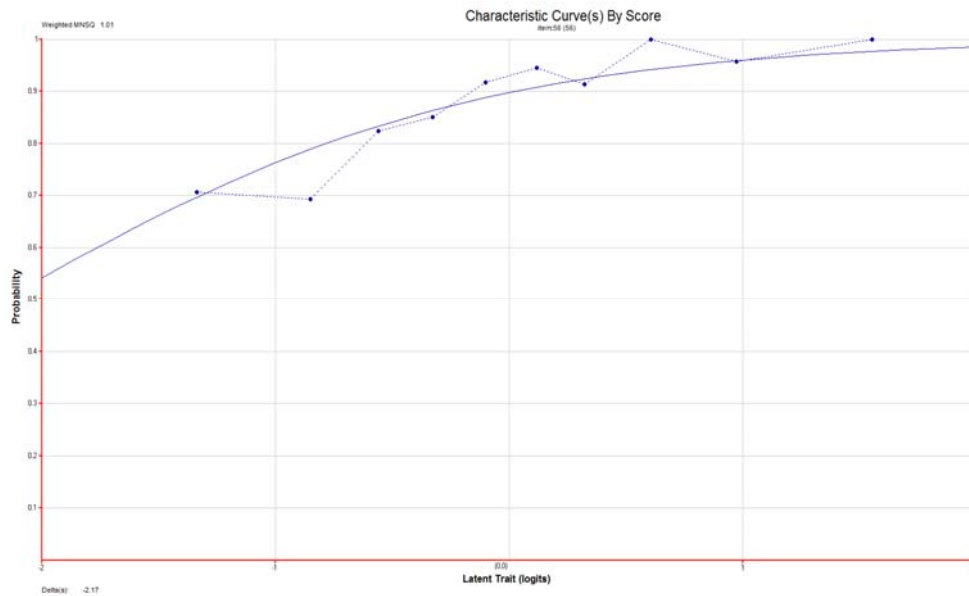


Abbildung 29 – ICC Item 60 (CBA)

Wie an Hand der Abbildungen zu sehen ist, sind insbesondere die Kurven der leichten Items (6, 59, 60) sehr flach, das bedeutet, dass die Items insgesamt wenig trennscharf sind und wie zu erwarten ist sowohl von sehr fähigen als auch von weniger kompetenten Probanden gleich wahrscheinlich beantwortet werden.

Item 49 und 64 hingegen nähern sich bezüglich der Abbildung der Trennschärfe der „idealen“-Form an:

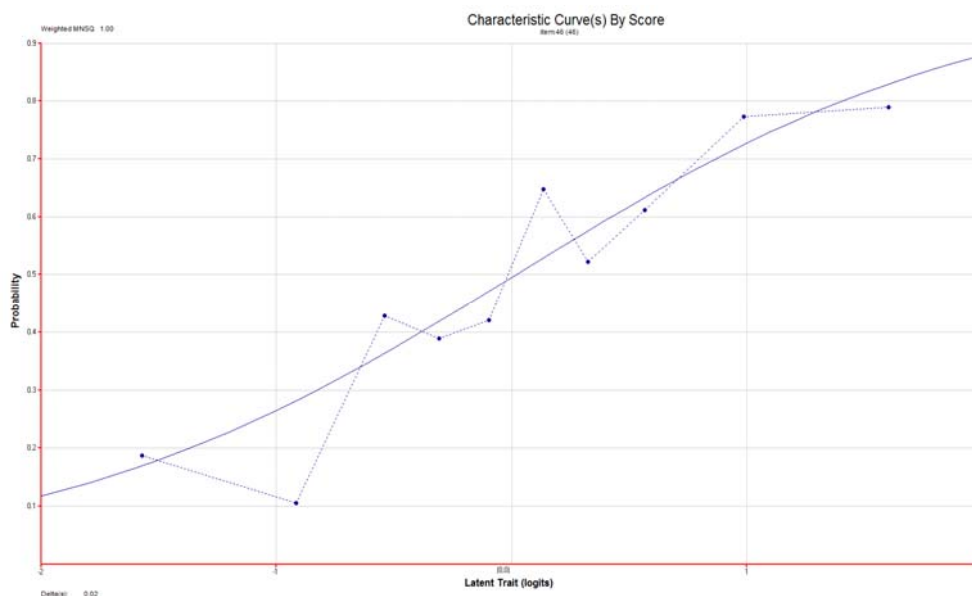


Abbildung 30 – ICC Item 49 (CBA)

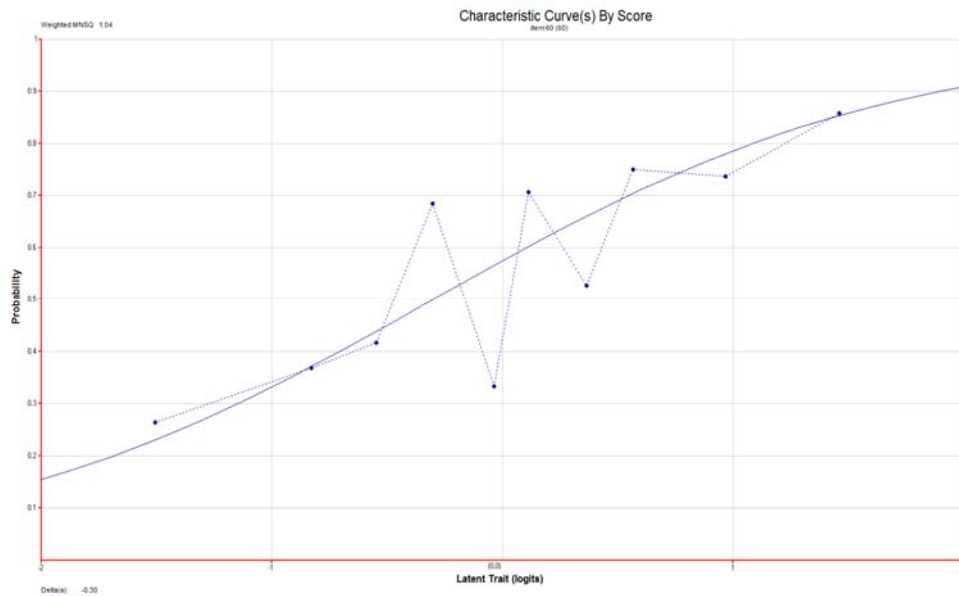


Abbildung 31 – ICC Item 64 (CBA)

Das schwerste Item (62) wiederum zeigt eine verhältnismäßig „steile“ bzw. nach unten gebogene Form, was bedeutet, dass dieses Item tendenziell sowohl von kompetenten als auch von weniger fähigen Personen mit geringerer Wahrscheinlichkeit beantwortet werden kann. Item 62 ist dabei noch in dem Rahmen, dass zwischen sehr fähigen und fähigen gut unterschieden werden kann.

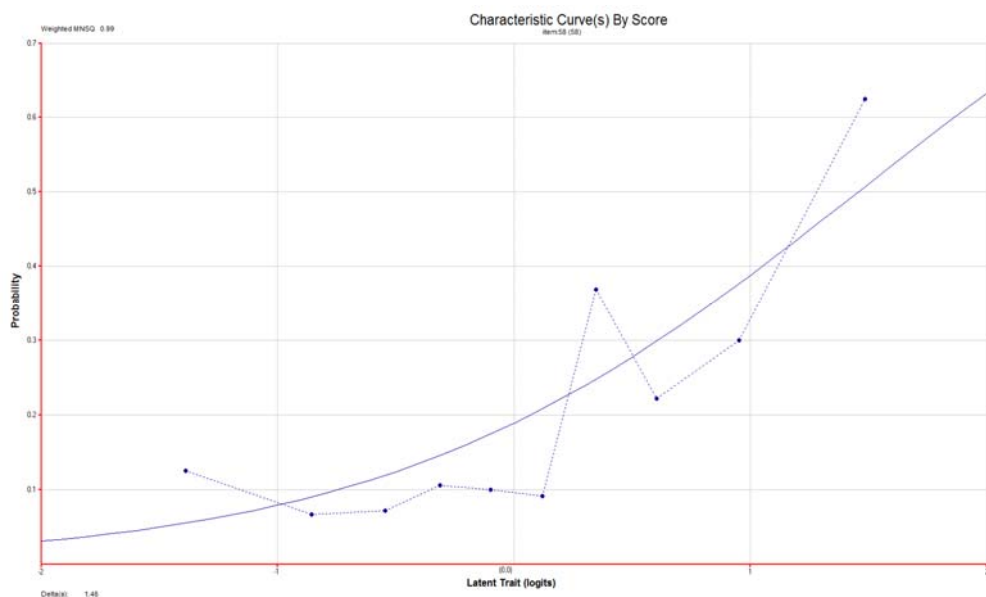


Abbildung 32 – ICC Item 62 (CBA)

Bezüglich des Fits ist bei Item 6, Item 49, Item 60 und Item 64 gut zu erkennen, dass die beobachtete ICC (durchgezogene Linie) weitgehend der erwarteten (punktierter Linie) entspricht. Item 59 hingegen ist im einfachen (linken) Bereich minimal steiler als erwartet, das weist auf einen leichten Überfit hin, was bei einem extrem leichten Item nicht weiter überrascht. Item 62 weist eine ähnliche Charakteristik im oberen (rechten) Bereich auf, ist dafür aber im unteren Bereich flacher als erwartet.

Neben der ICC gibt die Item Information Function (IFF) visuellen Aufschluss darüber, in welchem Fähigkeitsgehalt der Informationsgehalt am höchsten ist. Die Informationen entsprechen im Wesentlichen denen der ICC⁹⁹ (Abbildungen siehe Anhang S. 260-261).

Die Resultate der Kalibrierung werden an Hand des Antwortverhaltens der Probanden ermittelt, dabei wird jedoch nur zwischen der richtigen und falschen Antwort unterschieden. Ein Blick auf das erweiterte Antwortverhalten (welche Antworten wurden mit welcher Häufigkeit gewählt) kann weitere Einsichten auf die Items eröffnen, da so einerseits ermittelt werden kann, ob es eine besondere Tendenz zu einem Distraktor gibt – dies könnte ein Hinweis auf eine fehlerhafte Konstruktion sein, die Gleichverteilung der Antworten über alle Distraktoren (unabhängig von ihrer Bewertung als richtig oder falsch) ist ein Hinweis auf eine höhere Schwierigkeit der Frage und gleichzeitig spricht dies für eine gute Distraktorenkonstruktion.

Zu nächst zu Item 6 (vgl. Abbildung 15 auf S. 153):

Tabelle 8 – Antwortauswahl Item 6 (CBA)

	Häufigkeit	Prozent
Antwortmöglichkeit 1	14	7,3
Antwortmöglichkeit 2	7	3,6
Gültig Antwortmöglichkeit 3	3	1,6
Antwortmöglichkeit 4	168	87,5
Gesamt	192	100,0

Tabelle 9 – Einordnung der Antworten für Item 6 (CBA)

	Häufigkeit	Prozent
falsch	24	12,5
Gültig korrekt	168	87,5
Gesamt	192	100,0

Insgesamt antworteten 192 Personen auf dieses Item, davon 168 (87,5%) richtig und damit nur 24 (12,5%) falsch. Bezüglich der Distraktoren lässt sich bei dieser geringen Zahl der falschen Antworten keine interpretierbare Tendenz erkennen. Die hohe Tendenz zur richtigen Antwort zeigt die Konsistenz der Ergebnisse bezüglich der „Einfachheit“ des Items, da die Schwierigkeit über den Anteil richtiger Antworten geschätzt wird. Daher wird im Folgenden nur noch die Verteilung des Antwortverhaltens bezüglich der Distraktoren auf Hinweise auf Konstruktionsmängel o.ä. untersucht.

⁹⁹ Die approximierten Estimates geben den Bereich der Kurve (bzw. exakt den Höhepunkt der Funktion) an, in dem das Item seine beste Informationsleistung erbringt. Bei den leichten Items sind diese entsprechend der Estimate Werte links von 0 verschieden - respektive beim schwierigen Item rechts. Daran ist ebenfalls erkennbar, dass bei Items im mittleren Schwierigkeitsbereich entsprechend gute Werte für den Nullpunkt erreichen

Item 49 (vgl. Abbildung 17 auf S. 155) wurde entsprechend von 194 Probanden beantwortet. Im Gegensatz zu Item 6 antworteten hier nur 97 Personen korrekt – was exakt 50% entspricht. Von den falschen Antworten wies die Antwortmöglichkeit 4 eine nennenswert größere Präferenz innerhalb der Stichprobe auf (46 Personen bzw. 23,7%). Antwortmöglichkeit 2 hingegen wurde nur von 7,2% der Stichprobe gewählt. Dies weist auf ein nicht optimal konstruiertes Item hin, da dieser Distraktor als sehr unwahrscheinlich und somit als wenig geeignet bezeichnet werden muss.

Tabelle 10 – Antwortauswahl Item 49 (CBA)

	Häufigkeit	Prozent
Gültig	Antwortmöglichkeit 1	37
	Antwortmöglichkeit 2	14
	Antwortmöglichkeit 3	97
	Antwortmöglichkeit 4	46
	Gesamt	194
		100,0

Tabelle 11 – Einordnung der Antworten für Item 49 (CBA)

	Häufigkeit	Prozent
Gültig	falsch	97
	korrekt	97
	Gesamt	194
		100,0

Das scheint ungewöhnlich, da die beiden Informationen im Text keine lokale Kohärenz besitzen oder in irgendeiner Weise (bspw. durch Konnektoren oder Deiktika) direkt verbunden sind. Auch ist es nicht die erste in den Distraktoren wiederholte Information, die im Text erscheint – das wäre Antwortmöglichkeit 1. Diese belegt jedoch in der Antwortpräferenz nur Platz 3. Es mutet müßig an, weiter über das Antwortverhalten zu spekulieren, da diesbezüglich keine genaueren Daten existieren, wenngleich die Tatsache zumindest bemerkenswert scheint. Das dargestellte 50:50-Verhältnis illustriert sehr gut, dass das Item perfekt im mittleren Schwierigkeitsbereich liegt.

Ähnlich eindeutige Werte wie Item 6 weist Item 59 (vgl. Abbildung 14, S. 152) auf. Es wurde von 169 Probanden bearbeitet von denen 161 (95,3%) richtig antworteten. Hier wurden keine Distraktoren präsentiert, stattdessen gehört diese Item zur Klasse der offenen Antwortformate. Die Klassifikation der Antwortmöglichkeiten als richtig oder falsch liegt prinzipiell im Ermessen des jeweiligen Bewerter. So wurde hier entschieden bei allen offenen Items Rechtschreib-Fehler zu ignorieren, d.h. auch falsche Schreibweisen als richtig zu werten, da sie in diesem Kontext nicht ausschlaggebend waren (sofern nicht sinnentstellend). Dies hat selbstverständlich einen Einfluss auf die Schwierigkeit des Items in der Form das es leichter zu beantworten wird. Dafür sprach, dass somit Flüchtigkeits- und Tippfehler von der Beurteilung ausgeschlossen werden

können. An Hand der alternativ gegebenen Antworten lässt sich weiterhin feststellen, dass keine besonders häufig vorkommt.

Für Item 60 (vgl. Abbildung 12 auf S. 150) gilt, dass es von 202 Personen beantwortet wurde, davon antworteten 177, d.h. 87,6%, korrekt. Die Auffälligkeit bezüglich der Distraktoren bei diesem Item ist, dass kein einziger Proband Antwortmöglichkeit 4 wählte.

Tabelle 12 – Antworten Item 60 (CBA)

	Häufigkeit	Prozent
Antwortmöglichkeit 1	177	87,6
Gültig Antwortmöglichkeit 2	15	7,4
Antwortmöglichkeit 3	10	5,0
Gesamt	202	100,0

Tabelle 13 – Einordnung der Antworten für Item 60 (CBA)

	Häufigkeit	Prozent
falsch	25	12,4
Gültig korrekt	177	87,6
Gesamt	202	100,0

Es muss an dieser Stelle kritisch festgehalten werden, dass die Konstruktion dieses Distraktors als suboptimal zu bewerten ist. Dennoch ist auf Grund der vorliegenden mehr als eindeutigen Tendenz nicht davon auszugehen, dass ein wahrscheinlicherer Distraktor das Antwortverhalten massiv beeinflusst hätte.

Item 62 hingegen weist ähnliche Werte wie Item 49 auf und muss daher hier nicht weiter betrachtet werden.

Abschließend Item 64 (vgl. Abbildung 19 S. 156): Es wurde von 185 Probanden beantwortet – es beantworteten 107 Personen (57,8%) richtig. Auffällig ist hier bezüglich des Antwortverhaltens, dass es neben der Tendenz zur richtigen Antwort eine gewisse Präferenz (44 Personen; 23,8%) für Antwort 2 besteht und diese somit aus den falschen Antwortmöglichkeiten heraussticht.

Tabelle 14 – Antworten Item 64 (CBA)

	Häufigkeit	Prozent
Antwortmöglichkeit 1	18	9,7
Antwortmöglichkeit 2	44	23,8
Gültig Antwortmöglichkeit 3	107	57,8
Antwortmöglichkeit 4	16	8,6
Gesamt	185	100,0

Tabelle 15 – Einordnung der Antworten für Item 64 (CBA)

	Häufigkeit	Prozent
falsch	78	42,2
Gültig korrekt	107	57,8
Gesamt	185	100,0

Dies könnte sich möglicherweise dadurch erklären lassen, dass die Phrase „des **auf der Vorfahrtstraße nachfolgenden** Fahrzeugs“ [keine Hervorhebung im Item] nicht beachtet worden ist, da dies der signifikante Unterschied zwischen Antwortmöglichkeit 2 und 3 ist – dies ist allerdings eine höchst spekulative, wenngleich plausible Annahme. Weitere dahingehende Untersuchungen (Think-aloud-Befragung, direkt an die Bearbeitung anschließendes Interview) scheinen besonders lohnend, um durch Fehler die kognitiven Prozesse genauer erfassen und verstehen zu können.

An dieser Stelle werden die Daten aufgeschlüsselt nach kognitiven Anforderungsbereichen und Präsentationsformaten des Stimulus betrachtet. Daraus ergeben sich die unten dargestellten Übersichten (vgl. Tabelle 16 und

Tabelle 17 auf S. 180). Dazu wurden, um einen ersten Eindruck zu gewinnen, pro Klasse die jeweiligen Mittelwerte berechnet. Augenscheinlich hat das Repräsentationsformat einen minimalen Einfluss auf die Schwierigkeit der Items, da die Estimate-Werte (-.3 | -.27 | -.67) der depiktionalen Items etwas niedriger, somit die Items leichter, sind. Die kognitiven Anforderungsbereiche weisen einen starken Unterschied zwischen der Stufe „Identifizieren“ und den Stufen „Integrieren“ und „Generieren“ auf. „Identifizieren“ Items scheinen deutlich wahrscheinlicher (-1.08) richtig beantwortet zu werden als Items der anderen Stufen (-.18 | -.056). Bei dieser Betrachtung weisen „Generieren“ Items erwartungskonform eine geringere Wahrscheinlichkeit auf beantwortet zu werden, als „Integrieren“ Items – wenngleich der Unterschied nicht so groß wie zuvor ist. Bezüglich der Bearbeitungszeit kann festgestellt werden, dass deskriptionale Items die höchste Bearbeitungszeit haben, jedoch weicht diese nicht bemerkenswert von den übrigen Zeiten ab.

Werden beide Perspektiven kombiniert so werden diese Werte wahrscheinlich überaus stark von der Menge der Items und starker Wechselwirkungen pro Kategorie beeinflusst, so dass eine weitere Interpretation an dieser Stelle auf Grund der Inhomogenität der Itemanzahl unterbleibt.

Tabelle 16 – CBA Werte aufgeschlüsselt nach Repräsentationsformaten

Deskriptional		Hybrid		Depiktional	
Estimate	-.304	Estimate	-.275	Estimate	-.67
Error	.176	Error	.18	Error	.19
WMNSQ	1.002	WMNSQ	.998	WMNSQ	.981
t	.04	t	.038	t	-.126
Cases	189	Cases	188,2	Cases	190,3
DC	.449	DC	.443	DC	.442
Time	00:01:52	Time	00:01:47	Time	00:01:17

Tabelle 17 – CBA Werte aufgeschlüsselt nach kognitiven Anforderungsbereichen

Identifizieren		Integrieren		Generieren	
Estimate	-1.080	Estimate	-.176	Estimate	-.056
Error	.196	Error	.179	Error	.174
WMNSQ	.977	WMNSQ	0.995	WMNSQ	1.004
t	-.205	t	-0.079	t	.18
Cases	188	Cases	191	Cases	187,8
DC	.45	DC	.442	DC	.44
Time	00:01:18	Time	00:01:42	Time	00:01:50

Nach dieser Anscheinsanalyse lässt sich statistisch differenzierter feststellen, dass sich die Unterschiede in den Mittelwerten der *Repräsentationsformate* nach Durchführung einer Varianzanalyse als nicht-signifikant erwiesen ($F=.522$; $p=.56$).

Die Mittelwertsunterschiede der *kognitiven Anforderungsbereiche* hingegen waren bei Inhomogenität der Varianzen hoch signifikant ($p=.002$; $F=7,1$). Der Unterschied von -.88 zwischen „identifizieren“ und „integrieren“ und der Unterschied von -1.11 zwischen „identifizieren“ und „generieren“ sind auf einem .05-Niveau signifikant. Der Unterschied zwischen „integrieren“ und „generieren“ ist wiederum nicht-signifikant.

Im Folgenden werden vier Exkurse dargeboten, die einerseits das Präsentationsmedium (Computer oder Papier) als Einflussfaktor in den Blick nehmen, andererseits die Items unter schwierigkeitsbestimmenden Parametern, welche aus der Textverständlichkeitsforschung abgeleitet sind analysieren und drittens ebendiese Parameter eine kritischen Reflexion unterziehen. Viertens erfolgt ein Vergleich zwischen den Ergebnissen der ad hoc und der post hoc Schwierigkeitsbestimmung der Items.

4.2.2. Exkurs I – Einflüsse des Präsentationsmediums

Die Kalibrierungsstudie wurde mit am PC präsentierten Items durchgeführt. Da hinreichende Hinweise (Hansen et al. 1978; Wright/Lickorish 1983; Gould/Grischkowsky 1984; Kretzschmar/Pleimling 2011) vorliegen, dass das Präsentationsmedium auf die beim Lesen ablaufenden Prozesse einen nicht unerheblichen Einfluss hat – wenngleich dieser subjektiv sein kann und nicht unumstritten ist –, scheint der Versuch der Berücksichtigung des Präsentationsmediums angebracht.

Erkenntnisperspektive und Untersuchungsdesign

Die *zentrale Fragestellung* in diesem Kapitel ist: Gibt es bei der Bearbeitung der Items seitens der Probanden bezüglich der Schwierigkeit bedeutsame Unterschiede zwischen der Beantwortung der Items am PC gegenüber der papierbasierten Beantwortung? Dabei sollen auch testpraktische Unterschiede von papier-¹⁰⁰ und computeradministriertem Testmaterial berücksichtigt werden.

Dazu werden die Ergebnisse der zweiten Erhebungswelle – der *Pilotierungsstudie*, in der im Wesentlichen der adaptive Algorithmus getestet wurde, herangezogen. Im Rahmen dieser Erhebung wurde einem Teil der Stichprobe ausgewählte Items (Anzahl für Lesen: 33) in Form eines Papiertesthefts vorgelegt. Die dort gemachten Angaben wurden anschließend ebenfalls computerisiert, bereinigt, sowie bezüglich ihrer Plausibilität geprüft und anschließend betreffend der Schwierigkeit kalibriert. Diese Ergebnisse werden im Folgenden vorgestellt und mit den Ergebnissen der CBA Kalibrierung verglichen. Sowohl die CBA als auch die PP getesteten Probanden bekamen in der Pilotierung nur noch eine Domäne (Mathe, Lesen oder Nawi) vorgelegt. Weiterhin wurden in beiden Testformaten zusätzliche Informationen bspw. zum soziodemographischen Hintergrund und zur Motivation erhoben.

Auf Grund der begrenzten Möglichkeiten, die PP-basiertes Testen gegenüber CBA-basiertem bietet, mussten beim PP Test einige Einschränkungen gemacht werden:

Aus pragmatischen Gründe wurde kein Testheftdesign benutzt, so dass es – bis auf die jeweiligen Entscheidungen der Probanden zu blättern – eine lineare Reihenfolge der Items gibt, daher wird im Folgenden auch die jeweilige Positionsnummer des Items mitangegeben. Auch konnten auf Grund der Bearbeitungsdauer wie oben ausgeführt nicht alle konstruierten und nach der Kalibrierung beibehaltenen Leseitems in den Test aufgenommen werden, daher wurde eine inhaltliche Auswahl von 33 Items getroffen, wobei auf die Abbildung der Repräsentationsformate und Anforderungsbereiche geachtet wurde.

Des Weiteren konnten aus testökonomischen Gründen keine Bearbeitungszeiten erhoben werden und es musste aus praktikablen Gründen auf eine 1:1 Darstellung der Items wie in der CBA-Testung verzichtet werden¹⁰¹. Anzumerken ist weiterhin, dass die Ergebnisse auf Grund der

¹⁰⁰ im Folgenden: kurz PP für Paper-and-Pencil

¹⁰¹ Wegen des Gewichts der Testhefte und aus Kostengründen wurden in den PP Testheften 2 PowerPoint-Folien auf einer DIN-A4 Seite gedruckt.

kleineren Stichprobe trotzdem die gleiche Aussagekraft wie die der CBA Kalibrierung haben, da die Testhefte so verteilt wurden, dass jedes Item 179 Personen vorgelegt wurde. Damit wurde die Items theoretisch gegenüber die Kalibrierung (durchschnittlich 189 Personen pro Item) nur von geringfügig weniger Probanden bearbeitet (ausgenommen Missings durch Positionseffekte auf Grund des linearen Testverlaufs).

Pilotierung inkl. Paper-Pencil-Studie

Zur besseren Übersicht werden die drei Stichproben (Kalibrierung, Pilotierung CAT, Pilotierung PP) hier tabellarisch gegenüber gestellt.

Tabelle 18 - Stichproben (Kalibrierung, Pilotierung CAT, Pilotierung PP)

	Kalibrierung	Pilotierung–CAT	Pilotierung–PP
N	1.632	1.116	582
Geschlecht	46,3% ♀ 52,6% ♂ k. A.: 1,2%	37,5% ♀ 61,9% ♂ k. A.: 0,5%	39% ♀ 61% ♂ k. A.: 0%
Hauptsächlich ge- sprochene Sprache	86,9% Deutsch	85,8% Deutsch	86% Deutsch
Qualifikationsniveau (formale allgemein- bildende Abschlüsse)	28,5% AHR 62,1% FOR 7,2% HS k. A. 1,4%	29,6% AHR 61,6% FOR 7,1% HS k. A. 0,5%	19,5% AHR 66,5% FOR 11,9% HS k. A. 1,3%
Berufsbereich	35,8% gewerblich- technisch, 22,9% medizinisch- pflegend, 33,8% kaufmännisch- verwaltend, 5,5% sonstige, k. A. 2,5%	41,9% gewerblich- technisch, 15,9% medizinisch- pflegend, 33,7% kaufmännisch- verwaltend, 6,6% sonstige, k. A. 1,8%	42,2% gewerblich- technisch, 26,3% medizinisch- pflegend, 21,8% kaufmännisch- verwaltend, 7,8% sonstige, k. A. 1,9%

Es wurde sowohl im CAT als auch im PP pro Proband nur noch jeweils eine Domäne (Lesen, Mathe oder Naturwissenschaften) getestet. Hierbei ist als Ergebnis in aller Kürze festzuhalten, dass im CAT der adaptive Algorithmus erfolgreich eingesetzt wurde und sich in der Praxis zumindest bezüglich seiner grundsätzlichen Einsetzbarkeit bewähren konnte.

Die Ergebnisse auf Itemebene sind – analog zu Tabelle 6 (S. 169) – Tabelle 19 zu entnehmen, zusätzlich wurde, als weitere Information, die jeweilige Position des Items im Testheft hinzugefügt, jedoch wegen der Schwierigkeit diese zu erheben auf die Darstellung der Bearbeitungszeit verzichtet. Fehlende Werte wurden nicht berechnet, da diese zu keinem Zeitpunkt die Kalibrierung maßgeblich beeinflussten.

Tabelle 19 – Itemkennwerte PP

ID	Estimate	Error	WMNSQ	t	DC	Position
01	-3.792	0.462	1.00	0.1	0.51	17
04	-2.919	0.322	1.02	0.2	0.33	31
06	-2.843	0.310	0.97	-0.0	0.23	3
09	0.291	0.171	1.00	-0.1	0.40	4
10	-1.329	0.195	0.98	-0.2	0.26	30
11	1.023	0.187	0.97	-0.3	0.33	8
17	1.327	0.201	1.05	0.5	0.12	2
20	-1.207	0.189	0.98	-0.2	0.35	32
22	0.301	0.171	1.03	0.6	0.59	24
28	1.899	0.225	1.00	0.0	0.23	23
29	-1.743	0.220	0.97	-0.2	0.34	20
36	1.050	0.186	1.02	0.3	0.35	12
37	-2.304	0.256	1.00	0.1	0.52	33
38	0.633	0.177	1.01	0.2	0.37	14
39	1.198	0.192	1.02	0.2	0.34	5
40	0.361	0.176	0.99	-0.1	0.41	16
41	0.786	0.179	0.93	-1.1	0.27	19
42	-0.739	0.187	1.00	0.0	0.38	26
45	-0.283	0.178	1.05	0.9	0.45	28
46	-0.204	0.183	1.01	0.2	0.33	21
48	-0.052	0.181	1.03	0.5	0.13	1
50	2.203	0.274	1.00	0.0	0.31	15
52	0.379	0.188	1.09	1.6	0.32	10
53	-0.008	0.189	0.85	-2.8	0.26	7
54	-0.056	0.190	0.94	-1.1	0.17	29
62	0.477	0.199	1.00	0.1	0.40	6
63	-0.174	0.196	1.03	0.5	0.32	22
64	-0.634	0.204	0.94	-0.8	0.47	25
66	1.091	0.220	1.10	1.0	0.33	18
67	-0.856	0.210	1.07	0.9	0.35	27
69	-0.750	0.211	1.02	0.3	0.32	11
70	-2.396	0.331	0.92	-0.3	0.33	9

Werden ähnliche Betrachtungen wie beim CBA-basierten Testen angestellt, so lässt sich festhalten, dass der *WMNSQ* bei keinem Item einen Wert größer als 1.20 aufweist, was zeigt, dass alle Items einen zufriedenstellenden Fit bezüglich des Modells aufweisen, 7 Items besitzen einen exakten Fit, der maximale Wert beträgt 1.1 (Item 66) und der minimale 0.85 (Item 53). Da keiner der Werte kleiner als 0.75 ist, können alle Items als gut bewertet werden (vgl. Bond/Fox zitiert nach Winther 2010, S. 154).

Der *t-Wert* beträgt maximal 1.6 (Item 52) und ist damit sowohl gemäß liberaler Betrachtungen (<2) also auch nach konservativen (<1.96) für den PaperPencil-Test über alle Items hinweg als gut zu beurteilen.

Bezüglich der *Discrimination* lässt sich allerdings festhalten, dass diese in 9 Fällen Werte unter .3 annimmt (Item 6, 10, 17, 28, 41, 48, 53, 54, 72), von denen in 6 Fällen die Werte weniger als 0.25 betragen und damit nicht mehr im akzeptablen Bereich liegen, obwohl die Items einen guten Fit und einen guten t-Wert aufweisen sind diese kritisch zu bewerten und sollten ggf. durch andere Items, die zufriedenstellende Werte aufweisen, ersetzt werden, um die Testergebnisse nicht zu verzerren.

Zur den *Estimate-Werten* kann festgehalten werden, dass das Item mit der höchsten Lösungswahrscheinlichkeit (\triangleq das leichteste Item) einen Wert von -3.792 aufweist (Item 1) und das Item mit der geringsten Lösungswahrscheinlichkeit (\triangleq das schwierigste Item) einen Wert von 2.2 (Item 50) besitzt. Des Weiteren ist die durchschnittliche Schwierigkeit -.285, was knapp unter dem anzustrebenden 0 Punkt liegt.

Eine Aufstellung der ermittelten Werte (angegeben ist jeweils der Mittelwert) nach kognitiven Anforderungsbereichen und Präsentationsformaten wurde auch hier erstellt.

Diese Übersicht gestaltet sich wie folgt:

Tabelle 20 – PP Werte aufgeschlüsselt nach Repräsentationsformaten

Deskriptional		Hybrid		Depiktional	
Estimate	-.344	Estimate	.04	Estimate	-.566
Error	.243	Error	.196	Error	.219
WMNSQ	.996	WMNSQ	1.005	WMNSQ	.99
t	.078	t	.108	t	-.217
DC	.341	DC	.353	DC	.312

Tabelle 21 – PP Werte aufgeschlüsselt nach kognitiven Anforderungsbereichen

identifizieren		integrieren		generieren	
Estimate	-.722	Estimate	-.201	Estimate	-.042
Error	.212	Error	.224	Error	.214

WMNSQ	.987	WMNSQ	1.018	WMNSQ	.983
t	-.222	t	.35	t	-.233
DC	.318	DC	.328	DC	.355

Die Ergebnisse des PP-Tests sind auf Grund der geringen Anzahl der Items pro Kategorie mit Vorsicht zu betrachten. Daher seien an dieser Stelle nur die Ergebnisse diskutiert, die jeweils eine der Dimensionen betreffen (Anforderungsbereich oder Präsentationsformat) – die Detailangaben sind der Vollständigkeit halber trotzdem angegeben. Weitere Studien sollten durchgeführt werden, um diese zu verifizieren oder zu falsifizieren.

Das in den Ergebnissen der CBA-Testung gefundene Differential zwischen den *kognitiven Anforderungsbereichen* findet sich auch hier (-.722 | -.201 | -.042) – weniger stark, dafür aber konsequenter. Die „Identifizieren“-Items scheinen leichter (-.72) zu sein als die „Integrieren“-Items (-.20) und diese wiederum leichter (-.04) als die „Generieren“-Items zu sein. Bei der Bearbeitung der Items im Papierformat hat es den Anschein, dass das *Präsentationsformat* einen größeren Einfluss hat (-.344 | .04 | -.566). Die als „hybrid“ klassifizierten Items weisen eine geringere Lösungswahrscheinlichkeit (.04) auf als deskriptionale (-.34) oder depiktionale (-.56). Eine Varianzanalyse ergab für alle Betrachtungen jedoch nicht-signifikante Mittelwertsunterschiede, so dass die Unterschiede nicht weiter relevant sind. Das erscheint logisch, wenn sich vergegenwärtigt wird, dass bei diesen komplexen Items mit der geringen Anzahl an Items pro Kategorie einerseits große Varianzinhomogenitäten auftreten und andererseits wahrscheinlich Konfundierungen von schwierigkeitsbestimmenden Parametern innerhalb der Items auftreten.

Bezüglich des WMNSQ, der t-Werte und der Discrimination lässt sich lediglich Homogenität im guten Bereich feststellen, hier zeigt keine Klasse Auffälligkeiten.

Drei der oben vorgestellten Items wurden auch im PP-Test verwendet – Item 6, 62, 64. Deren Ergebnisse gestalten sich im PP-Test wie folgt:

Tabelle 22 – Itemkennwerte Items 6, 62 und 64 (PP)

ID	Estimate	Error	WMNSQ	t	DC	Position
06	-2.843	.31	.97	.00	.23	3
62	.477	.199	1	.10	.4	6
64	-.634	.204	.94	-.80	.47	25

Wie an Hand der *Estimates* zu sehen ist, ist das Items 6 deutlich dem leichten, Item 62 dem schweren und 64 dem mittleren Bereich zuzuordnen.

Die ICC und IFF der drei vorgestellten Items sind im Anhang zu finden (s. 263-265) und entsprechen bis auf Details, die später noch (vgl. Kap. 0) diskutiert werden, denen der CBA Kalibrierung.

Abschließend kann wiederum das detaillierte Antwortverhalten der Probanden einen Blick auf die Items eröffnen. Da nur eines Auffälligkeiten zeigt, wird auf die Darstellung der andern Items mit dem erwartungskonformen Antwortverhalten der Probanden verzichtet.

Item 64 wurde 179 Probanden vorgelegt, davon gaben 16 (8,9%) keine Antwort ab und 26 (14,5%) Personen haben (vermutlich aus Zeitgründen – das Item befand sich an Position 25/33) den Test vorher beendet.

Tabelle 23 – Antwortauswahl Item 64 (PP)

	Häufigkeit	Prozent
Antwortmöglichkeit 1	20	11,2
Antwortmöglichkeit 2	23	12,8
Antwortmöglichkeit 3	70	39,1
Gültig Antwortmöglichkeit 4	24	13,4
Vorher abgebrochen	26	14,5
keine Antwort abgegeben	16	8,9
Gesamt	179	100,0

Tabelle 24 – Einordnung der Antworten für Item 64 (PP)

	Häufigkeit	Prozent
falsch	67	37,4
korrekt	70	39,1
Gültig Vorher abgebrochen	26	14,5
keine Antwort abgegeben	16	8,9
Gesamt	179	100,0

Hier zeigt sich die Problematik eines linearen Testverlaufs. Tendenziell neigen Probanden dazu die Items in der vorgelegten Reihenfolge zu bearbeiten. So kommt es zu der Situation, dass die weiter hinten aufgeführten Items bei begrenzter Testzeit seltener beantwortet wurden als die weiter vorne stehenden (Item an Position 1: 178 gültige Antworten, Item an Position 33: 119 gültige Antworten). Das führte bei diesem Item zu fast einem Viertel an fehlenden Antworten und somit erfolgte die Berechnung der Schwierigkeit auf einer anderen Basis.

Vergleich CBA und PP – Administration und Itemkennwerte

Zur Analyse der wesentlichen Unterschiede zwischen computerbasiertem und PP-basiertem Testen erfolgt zunächst eine allgemeine Betrachtung auf Testebene, die die Konsequenzen beider Testmodi gegenüberstellt. Danach erfolgt eine spezielle Betrachtung auf Itemebene und eine Diskussion einzelner Items. Verglichen werden im Folgenden die Ergebnisse der Kalibrierung und der Pilotierung PP.

Werden zunächst also die Informationen bezüglich des gesamten Tests graphisch dargestellt (vgl. Abbildung 33 und Abbildung 34), so zeigt sich sehr deutlich, dass mit dem PP-Test die höheren Schwierigkeitsstufen nicht voll erfasst werden, und folglich die Trennschärfe deutlich geringer ist – ein Beispiel um das zu verdeutlichen: Eine Person mit dem Fähigkeitswert 0 erreicht im PP-Test (blau) einen Score von ca. 18, eine Person mit dem Fähigkeitswert 2 eine Score von 28; im CBA-Test (rot) hingegen erreicht eine Person mit dem Fähigkeitswert 0 einen Score von ca. 40 mit dem Fähigkeitswert 2 einen Score von ca. 60. Noch augenfälliger ist dieser Unterschied im Fähigkeitsbereich zwischen -1 und +1. Der CBA-Test kann hier wesentlich besser differenzieren. Des Gleichen spiegelt sich im Vergleich der TIFs – der CBA- Test liefert mehr Informationen.



Abbildung 33 – TCCs (CBA + PP)

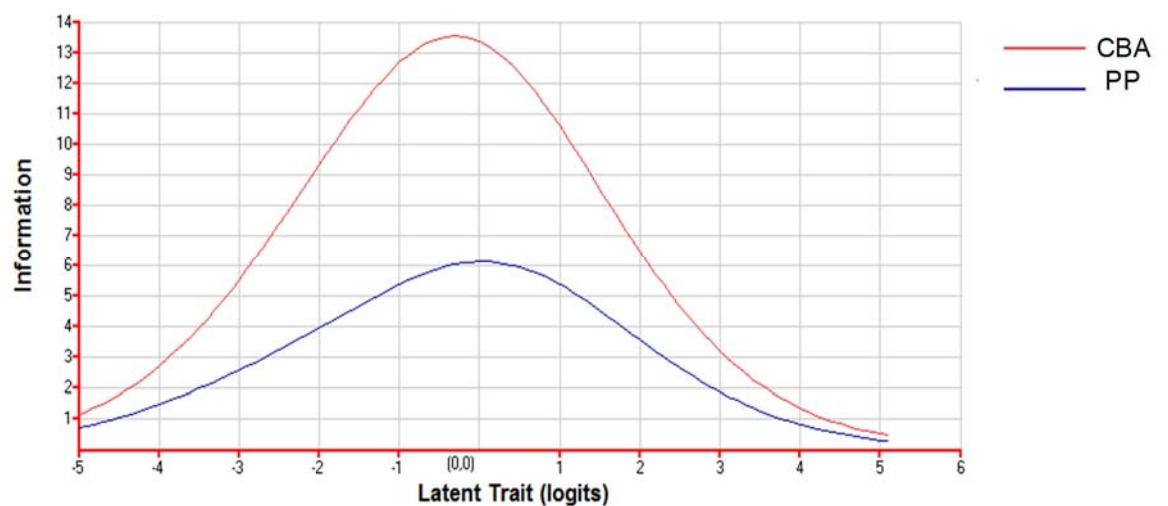


Abbildung 34 – TIF (CBA + PP)

Ursache dafür ist jedoch nicht direkt das Repräsentationsformat. Vermutlich liegen diese Gründe vornehmlich in der deutlich unterschiedlichen Itemanzahl – PP: 33 vs. CBA: 65. Das

zeigt bereits wenig überraschend, dass die Anzahl der Items einen großen Einfluss auf die Parameter Testinformation und auch Trennschärfe des Tests hat. Somit kann bereits an dieser Stelle festgehalten werden, dass einer der Vorteile computeradministrierten Testens die Möglichkeit ist ein komplexes Testheftdesign zu verwenden, in welchem eine größere Anzahl Items als Pool, aus dem gezogen wird, zur Verfügung steht. Insbesondere das adaptive Testen bietet hier mehr Möglichkeiten, da bereits mit weniger Testitems höhere Reliabilitätswerte erreicht werden können (vgl. dazu ausführlich Frey 2012 und Hinweise in Kapitel 5.1).

Um den Einfluss des Präsentationsmediums daher ansatzweise erfassen zu können ist es notwendig, an dieser Stelle nur die Items darzustellen, die in beiden Modi präsentiert wurden. Dazu sind die erhobenen Werte in einer Tabelle gegenüber gestellt (siehe Tabelle 46, S. 267).

Werden zunächst die *Estimate-Werte* global verglichen so fällt auf, dass es keine vollkommene Übereinstimmung gibt. Insgesamt weichen die mittleren Schwierigkeiten über den Gesamttest betrachtet nur unwesentlich (.12) nach unten ab, d.h. der PP-Test scheint global betrachtet minimal leichter – eine Korrelation der Estimates ergab jedoch nur einen nicht signifikanten Werte.

Eine Korrelation gibt die Stärke des Zusammenhangs zweier Größen an. Dabei wird die Stärke des linearen Zusammenhangs in einer singulären zwischen -1 und +1 liegenden Maßzahl angegeben. Ein positiver Wert deutet dabei auf einen positiven linearen Zusammenhang hin, ein negativer Wert kennzeichnet entsprechend einen negativen linearen Zusammenhang. Nochmals wichtig anzumerken ist, dass der Korrelationskoeffizient auf die Identifikation und Beschreibung linearer Zusammenhänge beschränkt ist. Daher kann auch bei einem Koeffizienten von null ein vollkommener, jedoch nicht linearer Zusammenhang zwischen den betrachteten Variablen vorliegen. Wenn der p-Wert Werte größer .05 annimmt, ist die Korrelation¹⁰² nicht signifikant. Somit ist das Ergebnis wahrscheinlich zufällig zustande gekommen und darf nicht interpretiert werden (vgl. Brosius 1998, S. 497).

Auf Item Ebene gibt es im Gegensatz dazu zum Teil große Differenzen. Besonders klein (.030) ist der Unterschied bei Item 20, kleiner als .4 ist er immerhin bei weiteren 6 Items (38, 69, 10, 36, 40 und 64). Besonders größere Differenzen zeigen sich bei Item 17 (3.85) und Item 1 (4.1), d.h. diese Items kehren ihre Polarität (leicht ↔ schwer) quasi um. Während Item 17 CBA-basiert eines der leichtesten Items war, wurde es bei PP-basierter Testung das dritt schwerste (ausführliche Diskussion siehe S. 191). Genau umgekehrt verhält es sich bei Item 1. Während es CBA-basiert im soliden Mittelfeld verortet ist, wird es im PP-Test zu dem mit Abstand leichtesten Item (ausführliche Diskussion siehe S. 190). Hierbei lässt sich die Erklärung jedoch nicht ausschließlich über das Präsentationsmedium finden – Motivations- und Konzentrationseffekte können ebenfalls eine Rolle spielen.

Bezüglich des *WMNSQ* und des *t-Werts* gibt es keine derartigen Unterschiede, die evident einen der Testmodi präferieren lassen.

¹⁰² gilt auch für andere statistische Kennwerte

Bezüglich der *Discrimination* ist global betrachtet festzustellen, dass die Differenz der Mittelwerte „nur“ .098 beträgt. Der kleinste Unterschied (.08) betrifft Item 1 und 64, der größte (.31) Item 54 und Item 39. Dennoch wird hier deutlich, dass ein globaler Vergleich nicht ausreichend ist, denn die in der CBA-basierten Testungen noch insgesamt zufriedenstellenden Werte, sind bei der PP-Testung zum Teil problematisch. So unterscheidet sich Item 17 um .3 und fällt damit von .42 auf .12 und damit deutlich unter die Grenze von .3. Ähnliches betrifft die Items 6, 10, 28, 41, 54 und 72. Das auch im CBA-Test bereits als kritisch beleuchtete Item 48 weist im PP-Test noch schlechtere Werte von nur .12 auf, was einer Differenz von .16 entspricht. Insgesamt ist die Trennschärfe beim PP-Test als schlechter zu bewerten – bei einem Mittelwerte von .329 statt wie CBA-basiert .427.

Zum Vergleich der Ergebnisse des CBA- und des PP-Tests bezüglich der kognitiven Anforderungsbereiche und der Repräsentationsformate ist es wie ausgeführt sinnvoll nur die Items zu vergleichen, die in beiden Testformen gelaufen sind.¹⁰³ In Tabelle 25 und

Tabelle 26 sind die entsprechenden Randsummen Werte für Anforderungsbereiche und Repräsentationsformate gegenübergestellt.

Tabelle 25 – Mittelwert der Estimates (PP und CBA_PP-Items) nach Repräsentationsformaten

Deskriptional		Hybrid		Depiktional	
CBA	PP	CBA	PP	CBA	PP
-.10	-.60	-.16	.22	.02	-.44

Tabelle 26 – Mittelwert der Estimates (PP und CBA_PP-Items) nach Anforderungsbereichen

Identifizieren		Integrieren		Generieren	
CBA	PP	CBA	PP	CBA	PP
-1.14	-.84	.12	-.49	.31	.51

Es zeigen sich zwar Unterschiede bezüglich der Mittelwerte, diese lassen jedoch keine Systematik erkennen, dass bspw. PP konsequent schwieriger oder leichter als CBA wäre. Bei den Repräsentationsformaten zeigt sich, dass bei im CBA die Werte von deskriptional zu hybrid leicht einfacher (.06) werden und zu depiktional wiederum schwerer (.18) als bei deskriptional und hybrid. Beim PP hingegen werden die Items von deskriptional zu hybrid deutlich schwerer (.82), und von hybrid zu depiktional wieder deutlich leichter (.66). Eine leichte Parallelität lässt sich bei den kognitiven Anforderungsbereichen erkennen. Sowohl beim CBA als auch beim PP werden diese von identifizieren zu integrieren und von integrieren zu generieren in unterschiedlichem Umfang schwerer.

¹⁰³ Um das Vorgehen nachvollziehbar zu machen, sind die CBA Mittelwerte der geschätzten Parameter in Tabelle 59 (S. 330) für die Items, die im PP gelaufen sind, abgetragen.

Werden die Schwierigkeitseinschätzungen für jede Kategorie jeweils korreliert, so ergeben sich nur für das Repräsentationsformat „deskriptional“ signifikante Zusammenhänge ($r = .79$, $p = .01$) der computerbasiert und den papierbasiert ermittelten Schwierigkeiten. Für die anderen Repräsentationsformate und die kognitiven Anforderungsbereiche ergeben sich lediglich nicht signifikante Werte. Auch eine Korrelation der Schwierigkeiten über alle Kategorien hinweg ergibt lediglich nicht signifikante Werte¹⁰⁴.

Auf Ebene der einzelnen Items lassen sich folgende Ergebnisse festhalten:

Der größte Unterschied (Estimate-Differenz: 4.1) zwischen der CBA Kalibrierung und den PP-Ergebnissen betrifft Item 1 (siehe Abbildung 35 und Abbildung 36). Anzumerken ist an dieser Stelle, dass in der Papierversion des Tests zwei Abweichungen zum CBA vorliegen. Einerseits werden die Items soweit nicht unbedingt nötig in s/w dargestellt, andererseits werden zwei PowerPoint-Folien auf einer DIN-A4 Seite dargestellt, so dass das Item den Probanden entsprechend Abbildung 35 und Abbildung 36 (S. 190) vorlag, während in der CBT-Variante geblättert werden musste.

Alkoholgehalt L0019321

Welches der folgenden Getränke enthält am meisten Alkohol? - 1 Pils (0,4l) - 1 Glas Weißwein, trocken (0,2l) - 1 Glas Sekt, trocken (0,1l) - 1 Weinbrand (0,01l)	Alkoholfreie Getränke	Energie (kcal/kJ) ¹
	Fruchtsaft, -nektar	50/210
	Fruchtsaftgetränk	50/210
	verdünnter Fruchtsaft	25/105
	Colagetränke	46/192
	Multi-Vitamin-Saft	47/196
	Karottensaft	27/113
	Tomatensaft	71/71
	Kaffee, ohne Milch und Zucker	2/8
	Tee, ohne Milch und Zucker	0/2
	Mineralwasser	0/0

Alkoholische Getränke	Energie (kcal/kJ) ¹	Alkohol (g) ¹
Bier (Pils)	42/176	3,9
Alkoholfreies Bier	28/117	0,2
Malzbier	52/218	1,0
Cidre	42/176	3,8
Rotwein, leicht	65/272	10,0
Weißwein, trocken	68/284	9,5
Sekt, trocken	76/318	10,0
Sherry, trocken	116/481	15,7
Fruchtsaftlikör	304/1271	24,0
Klarer, 32 Vol.-%	182/761	25,6
Weinbrand	234/978	33,0

¹pro 100 ml Getränk

105

Abbildung 35 – Item 1 Seite 1

¹⁰⁴ Dies liegt vermutlich an den kleinen Stichproben.

¹⁰⁵ Dieses Item wurde in Anlehnung an ein Item aus dem LTB³ entwickelt.

Welches der Getränke enthält am meisten Alkohol?

Bitte kreuzen Sie an.

- ☐ 1 Pils (0,4l)
- ☐ 1 Glas Weißwein, trocken (0,2l)
- ☐ 1 Glas Sekt (0,1l)
- ☐ 1 Weinbrand (0,01l)

Abbildung 36 – Item 1 Seite 2

Es gibt zwei wahrscheinliche Ursachen für die deutliche Vereinfachung des Items im PP gegenüber dem CBA: Erstens könnte es an der Darstellung auf einer Seite liegen, so dass die Probanden die verschiedenen Werte und Mengenangaben leichter vergleichen konnten; zweitens könnte es daran liegen, dass bei ohnehin Vorliegen der Frage auf Papier die Hemmschwelle und der Aufwand, schriftliche Nebenrechnungen durchzuführen, geringer ausfällt als bei einer computerbasierten Testung. Vereinzelt lassen sich derartige Hinweise auf Notizen der Probanden in den Testheften finden, wenngleich diese nicht aussagekräftig bzw. flächendeckend auftreten. Aus der Position des Items (Nr. 17 von 33) in der Mitte des Testhefts lassen sich keine Schlüsse ziehen, dass diese die Schwierigkeit derart positiv (im Sinne des leichter zu beantworten Seins) beeinflussen. Dergleichen gilt für mögliche Motivationseffekte.

Der zweit größte Unterschied (Estimate-Differenz: 3.85) zwischen der CBA Kalibrierung und den PP-Ergebnissen betrifft Item 17 (siehe Abbildung 37 und Abbildung 38).

Energiebedarf ermitteln L0179221

Tabelle 1 stellt den empfohlenen Energiebedarf für verschiedene Personen in Kilojoule (kJ) dar.

Bitte ermitteln Sie den empfohlenen Energiebedarf für Herrn Eder, einen 45-jährigen Lehrer.

Tabelle 1: EMPFOHLENER TÄGLICHER ENERGIEBEDARF FÜR ERWACHSENE

		MÄNNER	FRAUEN
Alter (Jahre)	Grad der Aktivität	Energiebedarf (kJ)	Energiebedarf (kJ)
Von 18 bis 29	Leicht	10660	8360
	Mittel	11080	8780
	Schwer	14420	9820
Von 30 bis 59	Leicht	10450	8570
	Mittel	12120	8990
	Schwer	14210	9790
60 und älter	Leicht	8780	7500
	Mittel	10240	7940
	Schwer	11910	8780

GRAD DER AKTIVITÄT NACH BERUFEN

Leicht:	Mittel:	Schwer:
Verkäufer	Lehrer	Bauarbeiter
Büroarbeit	Verkäufer (Außendienst)	Arbeiter
Hausfrau	Krankenschwester	Sportler

Abbildung 37 – Item 17 Seite 1

Energiebedarf ermitteln L0179221

Wie hoch ist der empfohlene tägliche Energiebedarf (in kJ) von Herrn Eder?

Bitte tragen Sie den Wert hier ein:

_____ kJ

Abbildung 38 – Item 17 Seite 2

Dieses Item wurde in der CBA Testung eindeutig als sehr leicht klassifiziert (Estimate: -2,525) – in der PP Testung ergab sich hingegen eine sehr geringe Lösungswahrscheinlichkeit (Estimate: 1,327), d.h. das Item gehörte zu den schwersten. Da das Item im Testheft an der zweiten Position steht, könnte es die Wirkung eines Positionseffekts sein (im Sinne, dass die Probanden

noch unsicher bezüglich der Antwortabgabe oder der gesamten Situation waren) oder auf materialbezogene Schwierigkeiten mit der Bearbeitung hindeuten. Hier scheint also die vorher getroffene Annahme, dass die gleichzeitige Darstellung von zwei ppt-Folien auf einer Seite zu einer Vereinfachung der Bearbeitung führt, widerlegt zu sein, da diese Darstellungsweise auch bei diesem Item erfolgte, hier jedoch scheinbar keinen Effekt hatte.

Daher lohnen sich Item 1 und Item 17 besonders für die Analysen in weiteren Untersuchungen – insbesondere Think-Aloud-Studien und eine Variation der Position könnten hier hilfreiche Hinweise für Prozesse oder Probleme bei der Bearbeitung liefern.

Zum Abschluss eine Analyse der 3 bereits weiter vorne vorgestellten Items, die sowohl im PP-Test als auch im CBA-Test gelaufen sind.

Tabelle 27 – Vergleich der Itemkennwerte CBA und PP für die Items 6, 62 und 64

ID	Estimate		Error		Fit		t		DC	
	CBA	PP	CBA	PP	CBA	PP	CBA	PP	CBA	PP
6	-2.10	-2.84	.23	.31	1.03	.97	.3	.00	.36	.23
62	1.43	.47	.19	.19	.99	1	-.1	.10	.41	.4
64	-.29	-.63	.16	.20	1.05	.94	.9	-.80	.39	.47

Zunächst lässt sich beobachten, dass alle drei Items leichter geworden sind.

Im Detail zeigt sich, dass Item 6 sich in seiner Schwierigkeit von CBA-Test zur PP-Version weiter in Richtung leicht verändert hat. Ein Grund dafür, dass dieses ohnehin schon leichte Item noch einfacher wird, könnte wiederum in der Darstellungsweise (2 Folien pro DIN-A4-Seite) liegen. Insgesamt scheint dies ein nicht zu ignorierender Faktor zu sein, da so u.a. Erinnerungseffekte beim Blättern nicht zum Tragen kommen, welchen Einfluss dies jedoch im Detail hat bleibt zu diesem Zeitpunkt ein Forschungsdesiderat.

Der Fit hat sich verbessert, die Trennschärfe deutlich verschlechtert. Es ergibt sich folgende Abbildung, wenn die beiden ICC übereinander gelegt werden. Wie u.a. zu erkennen ist, stellt sich bei der PP-Testung (blau) das Item einfacher da (wird bei geringer Fähigkeit mit höherer Wahrscheinlichkeit gelöst).

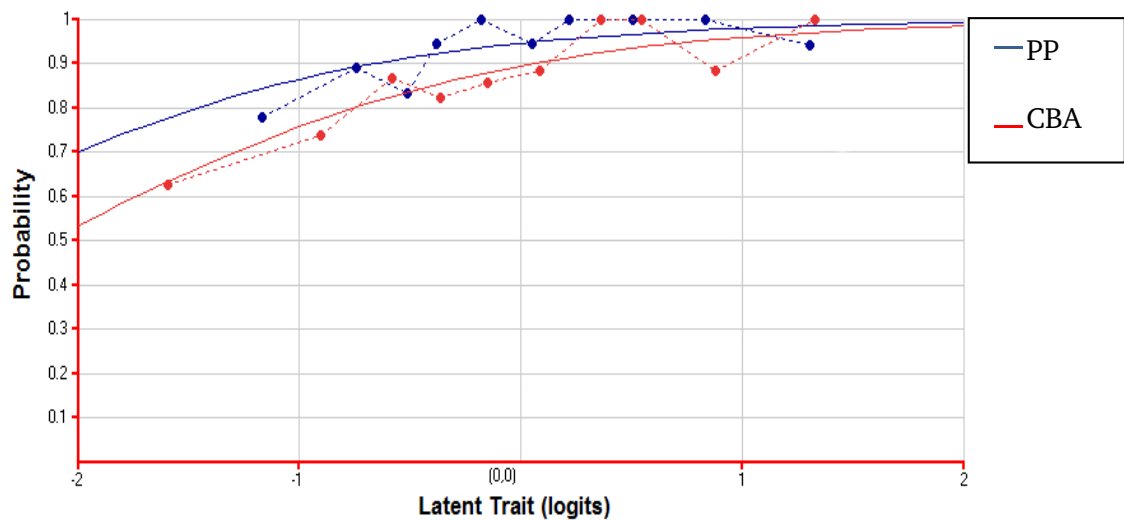


Abbildung 39 – ICCs Item 6 (CBA + PP)

Auch die Item Informationsfunktion unterstützt das bisher geschriebene. Den höchsten Informationsgehalt besitzt das Item wie aus den Estimates hervorgeht im PP (blau) weiter links auf der Fähigkeitsskala. Sie illustriert gut die Verschiebung, da Höhe und Breite insgesamt annähernd gleichbleiben.

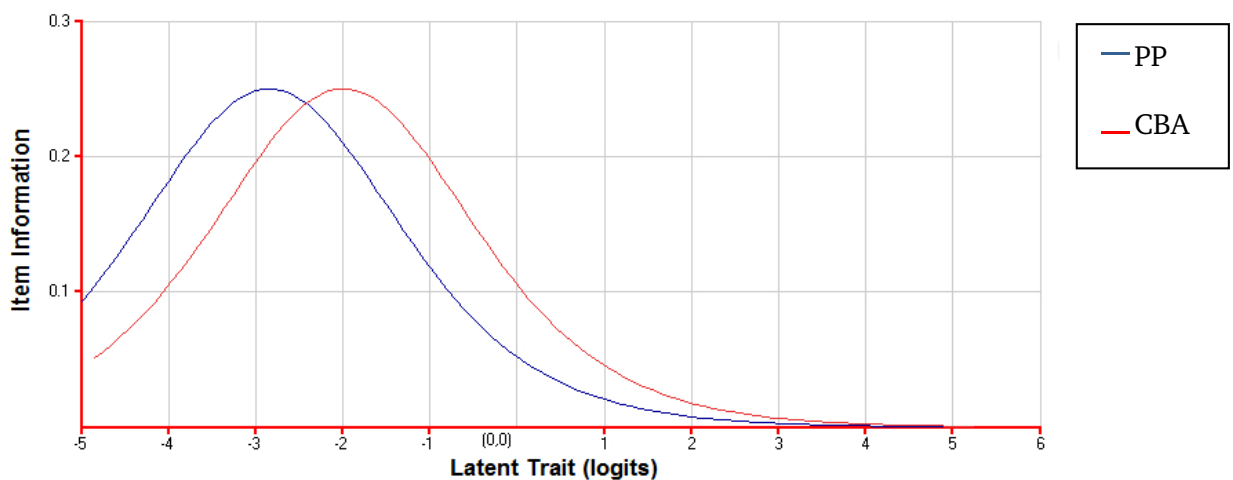


Abbildung 40 – IIFs Item 6 (CBA + PP)

Item 62 weist eine deutliche Vereinfachung um .95 auf. Auch der t-Wert ändert sich um .2. Der Fit und die Trennschärfe ändern sich nur minimal ($< .01$). In der Visualisierung stellen sich die Veränderungen – insbesondere, dass das Item einfacher zu lösen ist, wie folgt dar (siehe Abbildung 41; PP: blau | CBA: rot). Da die IIF dazu keine grundsätzlich neuen Erkenntnisse liefert wird diese nur der Vollständigkeit halber im Anhang aufgeführt.

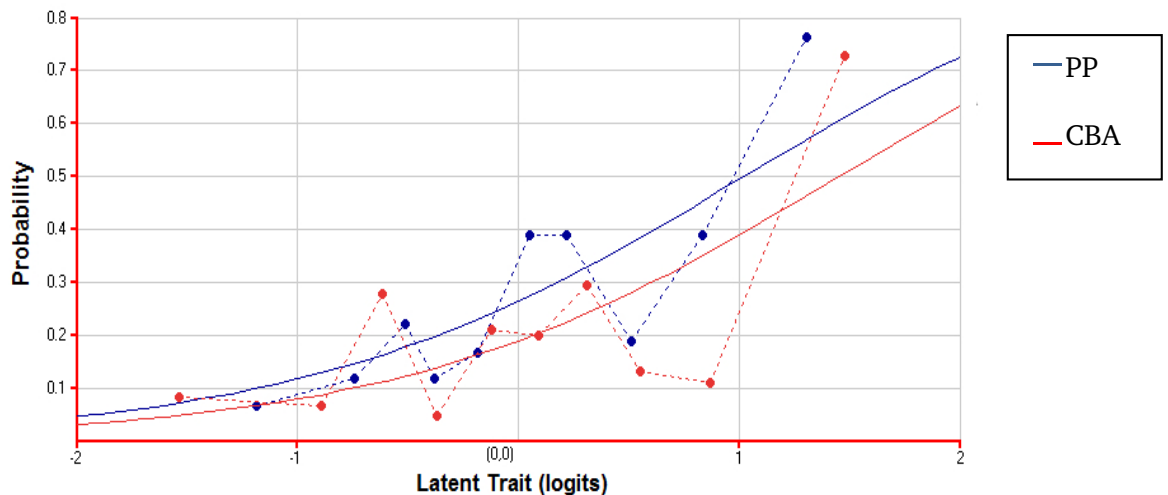


Abbildung 41 – ICCs Item 62 (CBA + PP)

Bei Item 64 gestaltet sich die Schwierigkeitsveränderung nur minimal (.3 – siehe Tabelle 27), ebenfalls so, dass das Item eine höhere Antwortwahrscheinlichkeit bei papierbasierter Testung aufweist. Daneben haben sich sowohl der der WMNSQ, um .1, als auch der t-Wert, um 1.7, , sowie die klassische Trennschärfe, um .08, verbessert. Die relativ ähnlichen Werte sind auch in der Illustration zu erkennen, die beiden Kurven sind annähernd gleich sehr (PP: blau | CBA: rot). Da die IIF dazu keine grundsätzlich neuen Erkenntnisse liefert wird diese nur der Vollständigkeit halber im Anhang aufgeführt.

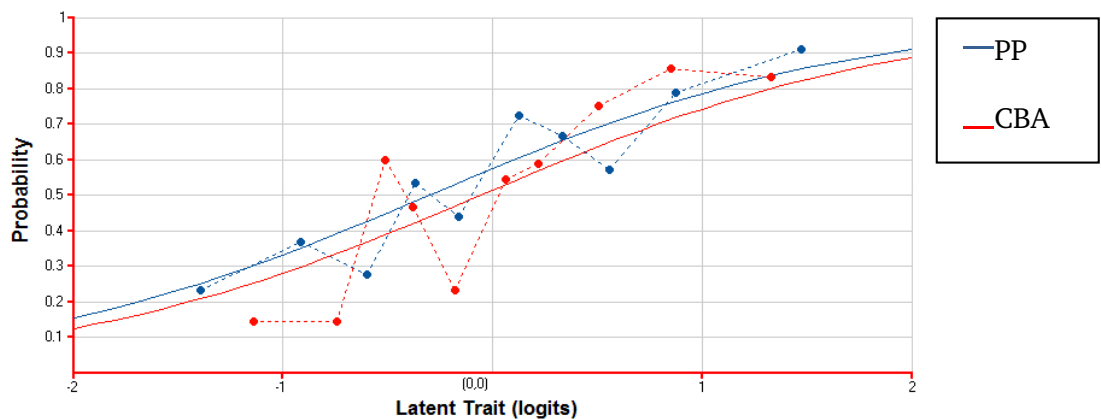


Abbildung 42 – ICCs Item 64 (CBA + PP)

Fasst man alle die präsentierten Ergebnisse – also über alle Items – zusammen, zeigt sich, dass der PP-Test insgesamt eine minimale Tendenz besitzt einfacher zu sein.

Unstrittig festgehalten werden kann, dass die beiden Tests nicht äquivalent sind. Dazu gibt es zu große Unterschiede bereits bezüglich der Durchführung (bspw. randomisierte vs. lineare Fragenpositionierung), was auf zum Teil starke motivationalen Beeinflussungen und Konzentrationseinflüsse geführt haben kann. Welcher Einfluss auf Positionseffekte zurückzuführen ist und welcher auf das Präsentationsmedium, kann nicht exakt bestimmt werden.

4.2.3. Exkurs II – Analyse der Itemschwierigkeit unter Aspekten der Verständlichkeitsforschung I

In Kapitel 3.1.5 wurde versucht an Hand der Textrezeptionsforschung herauszufinden, welche „Stellschrauben“ in einem Item möglicherweise angepasst werden können, um seine Beantwortung wahrscheinlicher, d.h. es selbst leichter zu machen und vice versa. Weitere Hinweise auf Variationsmöglichkeiten der Schwierigkeit wurden bereits in Kapitel 4.1.3 genannt. Die Ergebnisse waren jedoch nicht so eindeutig, dass direkte Konsequenzen für die Konstruktion daraus ableitbar wären. Daher wird hier post hoc versucht einzelne Parameter zu entwickeln und zu überprüfen, die zukünftig bei der Konstruktion von Leseitems bezüglich der Schwierigkeit hilfreich sein können.

Theoretischer Hintergrund: Schwierigkeitsbeeinflussende Parameter der Items – Ableitungen aus der Textverständlichkeitsforschung

In diesem Kapitel wird versucht konkrete schwierigkeitsbeeinflussende Parameter abgeleitet aus der Textverständlichkeitsforschung zu finden. Diese werden einer kritischen Untersuchung bezüglich Ihrer Objektivität (durch ein Ratingverfahren) und ihrer theoretischen Aussagekraft in Bezug auf die angenommene Itemschwierigkeit unterzogen. Abschließend werden die durch die Kriterien ermittelten Schwierigkeiten in Bezug zu den kognitiven Anforderungsbereichen gesetzt.

Eine Überprüfung der angenommenen Parameter bezüglich ihrer Aussagekraft über die Schwierigkeit der Items an Hand der Ergebnisse der Kalibrierung schließt sich im darauffolgenden Kapitel an (vgl. Kapitel 0).

Für eine a posteriori Betrachtung der Items können folgenden Merkmale herangezogen werden:

Ausgehend von dem Kriterium der *Oberflächenstruktur* kann der Umfang des gesamten Items, gemessen in PowerPoint Folien, berücksichtigt werden. Die Anzahl der Folien ist zumindest theoretisch bestimmend für die zu überblickende Menge an Informationen und möglichen Informationspositionen. Zusätzlich sind auch motivationale Einflüsse nicht auszuschließen. Zu bedenken bleibt, dass die reine Folienanzahl wenig über den konkreten Umfang des Inhalts aussagt, so dass bessere Indikatoren gefunden werden sollten. Gegen eine Zählung der Informationsmenge auf Wort-, Satz- oder Propositionsebene jedoch spricht die fehlende Äquivalenz zu depiktionalen Strukturen – hier muss ein weiteres Forschungsdesiderat festgehalten werden.

Ebenfalls auf Grund der fehlenden Entsprechung bei depiktionalen Formaten kann nicht auf *semantische sowie syntaktische Aspekte* – trotz ihrer sicherlich anzunehmenden Bedeutung für die Schwierigkeit – eingegangen werden.

Die *Aspekte der kognitiven Strukturierung, Prägnanz, Stimulanz/Verlebendigung oder literarisierende Merkmale* können nicht berücksichtigt werden, da diese zu sehr auf literarische oder didaktisierte Texte verweisen und im beruflichen Handlungskontexte diese Texte eher wenig vorkommen (vgl. Kapitel 2.1.2).

Die Notwendigkeit einer *Inferenzbildung* kann in drei möglichen Stufen berücksichtigt werden. Entweder müssen keine Inferenzen gebildet werden, es sind wenige (d.h. 1-2 Inferenzen) erforderlich oder drei bzw. mehr. Zu den Inferenzen gehören bspw. Generierung von Wortbedeutungen aus dem Kontext, Erkennen von Ursache und Wirkung, Konzepterkennung, Perspektivenwechsel oder die Herstellung logischer Verknüpfungen.

Die *Kohärenzdichte* kann an Hand verschiedener Subparameter abgebildet werden, welche nochmals in unterschiedliche Gruppen eingfasst sind. In der Gruppe der *Darbietung* des Items kann die *Art und Anzahl der jeweiligen Informationsquellen*, unter der Fragestellung, wo sich die zur Beantwortung nötigen Informationen befinden (innerhalb eines Dokuments, innerhalb mehrerer Dokumente des gleichen Repräsentationsmodus [Text/Text + Bild/Bild] oder innerhalb mehrerer Dokumente verschiedener Repräsentationsmodi [Text/Bild + Bild/Text]) erfasst werden.

In der Gruppe der *weiteren Merkmale* wird einerseits die Platzierung der zur Beantwortung nötigen Informationen, (was meint: befinden sie sich an einer zentralen oder einer versteckte Stelle, direkt in der Überschrift oder mitten im Textfluss, wurden sie gesondert hervorgehoben (fett, kursiv, andere Schriftfarbe) oder sind sie normal formatiert usw.), erfasst werden. Andererseits können in dieser Gruppe die *konkurrierenden Informationen* bspw. unwichtige Tabellenzellen, Abschnitte zu einem anderen Thema etc. im Text, berücksichtigt werden. Dabei wird nach ihrer Anzahl in den drei Abstufungen (keine, wenige (1-4), viele (>5)) gefragt.

Des Weiteren kann eine Gruppe mit dem Titel *Verarbeitungsprozesse* gebildet werden. Zwei weitere Parameter¹⁰⁶, die unter Textverarbeitungsprozesse gefasst werden, finden in der Textverstehensforschung so keine Entsprechung, da wie gezeigt hauptsächlich deskriptionale Texte betrachtet wurden: Die Berücksichtigung eines *Moduswechsels* (Schrift ↔ Bild) als Notwendigkeit zur Beantwortung der Frage. Dieses Kriterium kann dichotom konstruiert (ja/nein) werden. Es ist davon auszugehen, dass der kognitive Wechsel des Repräsentationsmodus keine einfache Aufgabe ist, da Fragen, die diese Transferleistung erfordern, sich nicht mehr über die propositionale Ebene beantworten lassen. Somit ist hierzu die Bildung eines Mentalen Modells zwingend notwendig (vgl. Kapitel 0). Der andere sich direkt hieraus ergebende Parameter ist die Frage nach der *Komplexität des* zu konstruierenden *Mentalen Modells*. Entweder ist kein Modell erforderlich oder es muss eines gebildet werden, dieses kann dann einfach, mittel oder komplex sein.

Im Rahmen einer Schwierigkeitseinschätzung noch zu berücksichtigen ist das *Antwortformat*. Rost (2004) führt aus, dass die wichtigste Unterscheidung von Antwortformaten die Trennung nach offenem oder geschlossenem Antwortformat ist. Bei einem offenem Antwortformat wird die Itemantwort von der getesteten Person selbst in einem (allgemein verständlichen) Zeichensystem formuliert wie z.B. in der Sprache, in Form von Zahlen, in Bildern.... Es bleibt anschlie-

¹⁰⁶ Je nach Betrachtungsweise könnte das Kriterium Inferenzbildung möglicherweise auch noch zu den Aspekten der Kohärenzdichte gezählt werden.

ßend dem Testleiter vorbehalten, diese gegebene Antwort in ein vorgefertigtes Kategoriensystem – hier: richtig/falsch – einzuordnen. Auch offene Antworten besitzen dabei ein Format, da vorgegeben wird, welche Art von Antwort die Person produzieren soll, bspw. ein Bild malen, einen Satz ergänzen, ein Muster fortsetzen, eine Zahlenreihe ergänzen oder eine Geschichte erzählen. Ein starkes Argument für die Bevorzugung dieses Formats insbesondere bei Leistungstests ist, dass freie Antworten ein Mittel sein können, um die Wahrscheinlichkeit einzuschränken, dass die richtige Antwort erraten wird (vgl. Rost 2004, S. 61). Somit ist deutlich, dass die Formulierung einer offenen Lösung als schwieriger gelten kann, als das Ankreuzen einer vorgefertigten Antwort. Was gegen diese Betrachtungsweise spricht wird im Folgenden noch ausgeführt.

Ein geschlossenes Antwortformat bietet gegenüber dem offenen eine Auswahl von Verhaltensalternativen an. Die Person braucht somit die Itemantwort nicht zu artikulieren, sondern hat einen begrenzten Antwortbereich zur Verfügung, aus welchem eine Auswahl zu treffen ist. Der Vorteil dieser Antwortformate liegt darin, dass der Prozess der Signierung¹⁰⁷ sich erübrigt. Gebundene Antwortformate haben nach Rost zunächst den Anschein einer höheren Objektivität und sind durchaus häufig objektiver, da die Auswertungsobjektivität sehr hoch ist. Die durch die vorgegebenen Antwortalternativen erzwungene Objektivität kann jedoch auch schnell negativ auf die Validität des Tests wirken, da die vorgegebenen Alternativen möglicherweise nicht alle Reaktionsmöglichkeiten umfassen, das Durchlesen der Alternativen das Antwortverhalten beeinflusst oder die vorgegebenen Antworten in ihrer Formulierung und Duktus nicht der natürlichen Reaktion der befragten Person entspricht (vgl. Rost 2004, S. 63).

Der Hauptvorteil geschlossener Formate besteht in der Auswertungsökonomie des Tests, d.h. derartige Tests sind quasi sofort, von ungeschulten Auswertern und mit Schablonen auswertbar und somit optimal bei großen Stichproben einsetzbar. Tests mit offenen Antworten können gemeinhin die gleiche Stufe an Objektivität haben, aber vermutlich verbunden mit einem höheren Aufwand (vgl. Rost 2004, S. 63).

Bei den vorgegebenen Antwortkategorien eines gebundenen Antwortformates ist zu beachten, ob die Kategorien disjunkt sind¹⁰⁸ und ob die Menge der vorgegebenen Kategorien exhaustiv ist¹⁰⁹. Bei Leistungstests sind nicht-exhaustive Formate sehr gebräuchlich, da sie meist gar nicht exhaustiv sein können. Denn Leistungsfragen haben eine potentiell unendlich große Anzahl denkbarer (insbesondere falscher) Itemantworten, aus der nur eine kleine Anzahl zur Auswahl offeriert werden kann – die richtige sollte selbstverständlich darunter sein (vgl. Rost 2004, S. 63f).

Die aus der großen Anzahl potentieller falscher Antworten ausgewählten Alternativen werden als Distraktoren bezeichnet. Diese haben die Funktion, die Identifikation der richtigen Antwort

¹⁰⁷ Einordnung der Itemantwort in Verhaltenskategorien

¹⁰⁸ D.h. einander ausschließen

¹⁰⁹ D.h. den Bereich aller Antwortmöglichkeiten ausschöpft

zu erschweren. Dafür besteht eine Notwendigkeit, da „der Lösungsprozess bei gebundenen Antwortformaten grundsätzlich ein anderer ist als bei freien Antworten“ (Rost 2004, S. 64). Dieser läuft bei vorgegebenen Antwortalternativen in der Regel in der Form, dass alle präsentierten Alternativen daraufhin geprüft werden, ob sie eine adäquate Itemantwort darstellen. Je kongruenter die Antwortkategorien sind, desto schwieriger gestaltet sich der Auswahlprozess für die Probanden. Bei Leistungstestitems wird dieser Auswahlprozess nicht nur durch die Ähnlichkeit der Antwortalternativen erschwert, sondern auch durch die Plausibilität der Distraktoren. Der zeitliche Aufwand zur Lösungsfindung kann approximativ nach Belieben durch das Angebot sehr schwieriger Distraktoren gesteigert werden (vgl. Rost 2004, S. 64f).

Während die Antwortkategorien bei Leistungstests prinzipiell nicht exhaustiv (s.o.) sein können, ist eine wichtige Anforderung, wenn nur eine Antwortalternative ausgewählt werden darf, dass die Alternativen disjunkt sind. Das ist erforderlich, damit die Probanden ihre Antwort exakt in einer der vorgegebenen Antwortkategorien identifizieren können (vgl. Rost 2004, S. 65).

Bei Leistungstests wird vielfach ein Kunstgriff benützt, um die Ratewahrscheinlichkeit weiter zu senken. Bei Auswahl von nur einer Kategorie aus n vorgegebenen Kategorien beträgt die Ratewahrscheinlichkeit nämlich $1/n$, also bei 5 Antwortalternativen 20%. Soll dagegen aus fünf Antwortalternativen zwei ausgewählt werden, verringert sich die Ratewahrscheinlichkeit bereits auf 10%, da die Anzahl der möglichen Zweierkombinationen aus fünf Elementen $(5 \cdot 4)/2 = 10$ beträgt. Die Ratewahrscheinlichkeit wird daher minimal, wenn die Anzahl richtiger Antworten nicht vorgegeben wird, sondern es den Probanden freigestellt wird, wie viele Alternativen für richtig gehalten werden (vgl. Rost 2004, S. 65f).

Auf Grund der projektbezogenen Vorgaben bezüglich der sofortigen computerisierten Auswertbarkeit waren längere offene Antworten, Partial-Credit-Modellierungen sowie Mehrfachauswahl Antworten nicht möglich.

Daher kann bei der Schwierigkeitseinschätzung betreffend der Distraktorengestaltung nur berücksichtigt werden, ob die Antworten auf das Item geschlossen oder offen erfolgen sowie die Anzahl plausibler Antwortalternativen.

Der *Vorwissens- bzw. Weltwissensbezug* kann berücksichtigt werden, allerdings anders als es die Textverstehensforschung nahelegt. Wenn ein Proband besonderes Weltwissen, d.h. Wissen über die Sachverhalte des Textinhalts, benötigt, um die Frage zu beantworten, so ist diese als schwerer zu beantworten als eine autark zu beantwortende Frage. Bei der Messung von Lesekompetenz kommt es im besonderem Maße darauf an, dass keine außerhalb der eigentlich Kompetenz liegenden Faktoren das Ergebnis beeinflussen und so die Messung verzerren, daher sind alle externen Faktoren nach Möglichkeit auszuschließen – in diesem Fall das Weltwissen. Das Textsortenwissen hingegen oder in diesem Kontext besser bezeichnet als *Text-/Grafikstrukturwissen* (Konvention wie etwas ausgedrückt wird, wie etwas dargestellt wird), kann als Parameter bei der Konstruktion berücksichtigt werden, da dies als ein Teil der Lesekompetenz verstanden werden kann. Beides kann dichotom (ja/nein) erfasst werden.

Der *inhaltliche Anspruch* des Stimulus kann bei diesen Kriterien nicht berücksichtigt werden – wie die inhaltlichen Anforderungen eines Textes mittels objektiver, d.h. subjektunabhängiger, Ratingparameter erfasst werden können ist nicht abschließend geklärt und kann daher hier nicht weiter operationalisiert werden.

Erkenntnisperspektive und Untersuchungsdesign

Die hier im Fokus stehende Frage ist:

Ist bei einer kriteriengeleiteten Schwierigkeitseinschätzung der konstruierten Items die Rater-reliabilität dahingehend zufriedenstellend, dass dieses Verfahren die Schwierigkeit verhältnismäßig objektiv, d.h. unabhängig von subjektiven Einschätzungen der Rater, abbildet?

Damit ist impliziert, dass die Parameter hinreichend eindeutig formuliert sein müssen, so dass sie von verschiedenen Personen korrespondierend identifiziert werden.

Um diese Frage untersuchen zu können, wird wie folgt vorgegangen: Zunächst werden zwei Ratern alle 65110 Items präsentiert. Beide Rater waren in das Projekt involviert und können daher als sehr gut bis gut geschult gelten¹¹¹. Der Ratingbogen mit den Parametern wurde am Computer vorgelegt¹¹² und wurde pro Item separat ausgefüllt. Die Ergebnisse konnten untereinander (sowohl rater- als auch itembezogen) während des Ratings nicht eingesehen werden. Abschließend wurde eine generelle Einschätzung der Itemschwierigkeit in den drei Abstufungen „einfach“, „mittel“, „schwer“ erbeten, so dass die Rater eine persönliche intuitive Einschätzung angeben konnten. Inwiefern diese mit den Ergebnissen der Summe der einzelnen Parameter übereinstimmt wird ebenfalls betrachtet. Diese Frage bezüglich der Einschätzung der Gesamtschwierigkeit des Items wurde gesondert von den anderen Fragen präsentiert, so dass diese Antworten nicht mehr vorlagen. Um zu überprüfen in wie weit die Ergebnisse von der Präsentation der Parameter beeinflusst werden können, wurde diese Frage sowohl vor als auch nach der Präsentation der Parameter gestellt.

Zur Überprüfung des Kategoriensystems werden zunächst Analysen hinsichtlich der Übereinstimmung zwischen den Ratern durchgeführt. Im ersten Schritt global über alle Items, danach itemspezifisch. Bei diesem zweiten Schritt werden die Itemschwierigkeiten an Hand eines Punkterasters dargestellt und in ein Niveaustufensystem eingebettet. Dann wird auch untersucht inwieweit einerseits die unterschiedlichen Schwierigkeitseinschätzungen der Rater korrelieren und in welchem Grad es eine Übereinstimmung zwischen der Schwierigkeitsberechnung via der Parameter und der intuitiven Einschätzung der Gesamtschwierigkeit gibt.

¹¹⁰ Auf Grund technischer Erfordernisse und geschlechtsbezogenen Diff-Effekten, die in der Kalibrierung ermittelt wurden (vgl. Kapitel 0), wurden 8 Items (ID: 22, 26, 30, 31, 44, 53, 65 und 71) ausgeschlossen und nur noch die übrigen geratet.

¹¹¹ Rater A war wissenschaftlicher Mitarbeiter im Projekt und hauptsächlich für die Konstruktion der Items verantwortlich, Rater B war wissenschaftliche Hilfskraft im Projekt und größtenteils mit der Korrektur und Anpassung der Items betraut

¹¹² Der Bogen ist in seiner Gesamtheit unter Kapitel 9 – S. 310 einzusehen.

Abschließend werden die Ratingergebnisse bezüglich der als Punkte ermittelten Schwierigkeiten in einer Tabelle aufgeschlüsselt nach kognitiven Anforderungsbereichen und Präsentationsformate abgetragen und diskutiert.

Ratingergebnisse

Zur Überprüfung des Kategoriensystems werden zunächst die Ergebnisse der Analysen hinsichtlich der Übereinstimmung zwischen den Ratern global über alle Items präsentiert. Untersucht wurde, in wie weit jedes theoretisch schwierigkeitsbestimmende Merkmal von den Ratern itembezogen gleich eingeschätzt wird.

Zur Bestimmung der Interoderreliabilität wird im Folgenden der verhältnismäßig robuste Holsti-Koeffizient ($=R$) verwendet. Der Holsti-Reliabilitätskoeffizient beruht auf der Logik des Paarvergleichs (vgl. Rössler 2005, S. 190). Dabei erfolgt „die Berechnung des Holsti-Koeffizienten ... zweistufig: Zuerst berechnet man alle möglichen Paarvergleiche und bestimmt anschließend den Durchschnitt aller möglichen Paarvergleiche“ (Raupp/Vogelsang 2009, S. XVI). Die von Merten (1983) in seiner Kritik am Holsti-Koeffizienten aufgezeigte Gefahr von Schein-Reliabilitäten infolge von Zufallsübereinstimmungen erscheint angesichts von 65 Merkmalen¹¹³ als im Rahmen dieser Arbeit zu vernachlässigende Größe. Mertens Plädoyer für die Berücksichtigung von Zufallsübereinstimmungen etwa mittels Cohens Kappa-Koeffizienten – bei dem die zufallskorrigierte Übereinstimmung zweier Gutachter bei der Kodierung mit einem vorgegebenen Kategoriensystem berechnet wird –, ist zwar ein berechtigter Hinweis, der allerdings vor allem für Kodierungen mit nur wenigen Kategorien gilt (vgl. Merten 1983, S. 304f).

Zusätzlich werden zur besseren Einordnung Prozentwerte angegeben, welche eine oberflächliche, aber schnelle Interpretation erlauben. Dass es dabei zum Teil zu irritierenden Abweichungen kommen kann, liegt an der oben ausgeführten Paarvergleichslogik, die dem Holsti-Koeffizienten zugrunde liegt. Dies betrifft vor allem Kategorien mit mehr als zwei Merkmalen. Prozentual kann hier nur stimmt überein vs. stimmt nicht überein geprüft werden, bei der Berechnung des Holstikoeffizienten hingegen, wird dabei auch die Größe der Abweichung berücksichtigt – wie weit liegen die Unterschiede auseinander?

Tabelle 28 – Raterreliabilitäten der schwierigkeitsbestimmenden Merkmale

Relevante Merkmal	Prozentuale Übereinstimmung	R
Seitenzahl	100%	1.00
Antwortformat	100%	1.00
Art + Anzahl der Quellen	86.2%	0.98
Inferenzbildung	55.4%	0.65
Irrelevante Informationen	47.7%	0.93
Mentalen Modells notwendig	52.3%	0.55

¹¹³ Bei einer Betrachtung der Parameter global über alle Items hinweg werden die Items zu den relevanten Merkmalen.

Moduswechsel notwendig	81.5%	0.85
Platzierung der relevanten Information	73.8%	0.86
bes. Sprachwissen notwendig	83.1%	0.92
Format des Stimulus	75.4%	0.77
bes. Weltwissen notwendig	93.8%	0.94

An Hand von Tabelle 28 ist zu ermitteln, dass die durchschnittliche prozentuale Übereinstimmung bei 77.2% und $R = .85$ liegt. Somit ist das Ergebnis diese Analyse durchaus als zufriedenstellend zu betrachten.

Auch ist eine Betrachtung von Details nicht uninteressant. Das Merkmal mit dem geringsten Wert nach Holsti ist die Frage nach der Notwendigkeit und Komplexität des Mentalen Modells – das Merkmal mit der geringsten prozentualen Übereinstimmung ist das nach der Anzahl der irrelevanten Informationen.

Auch wenn für diese Beobachtungen verschiedene Ursachen in Frage kommen, kann es eine gemeinsame Erklärung geben: Einerseits sind die beim verstehenden Lesen ablaufenden Prozesse zwar, wie bereits in Kapitel 3.1.4 ausführlich dargestellt, kognitiv höchst komplex, allerdings laufen viele davon insbesondere wenn es nicht zu Störungen im Prozess kommt automatisch und damit evtl. unbewusst ab. Dies könnte besonders bei geübten und kompetenten Lesern – wie bspw. Beschäftigten im akademischen Bereich – zu einer Verzerrung der Wahrnehmung führen. Fraglich ist also, ob die irrelevanten Informationen und die anspruchsvolle Konstruktion des mentalen Modells überhaupt bewusst wahrgenommen werden oder, ob noch vorbewusst die entsprechenden Informationen z.B. als irrelevant identifiziert und damit ausgeblendet werden, so dass auch bei einer Nachfrage die Berücksichtigungsnotwendigkeit dieser Daten marginalisiert wird.

Dies könnte evtl. zukünftig durch eine ausführlichere Raterschulung und Bewusstmachung dieser Problematik verbessert werden. Eine andere mögliche Erklärung ist ggf. die jeweils individuelle Varianz, da diese abhängig vom Vorwissen des Lesers bzw. Raters ist.

Dass sich besonders hohe Übereinstimmungen beim Antwortformat und der Seitenzahl ergeben, sollte nicht überraschen.

Positiv ist, dass große Übereinstimmung (bezogen auf die Prozentzahl – 93,8%) bezüglich der Frage nach der Notwendigkeit der Nutzung speziellen Weltwissens ergeben. Zum Großteil lautet hier die Einschätzung „Nein, kein besonderes Weltwissen notwendig“, so dass das Konstruktionsziel vom Weltwissen autarke Items zu entwickeln (vgl. O, S. 199) als erfüllt angesehen werden kann. Bezogen auf den Holsti-Koeffizienten liefert das Merkmal „Art und Anzahl der Quellen“ besonders gute Werte (.98) was ebenfalls zufriedenstellend ist, da die Frage, ob Quellen zum gleichen Repräsentationsmodus gehören oder eher doch verschiedenen entsprechen, zentral für die Theorie und die Anforderungseinordnungen ist.

Zur Überprüfung des Kategoriensystems werden nun die Ergebnisse der Analysen hinsichtlich der Übereinstimmung zwischen den Ratern *itemspezifisch* dargestellt, d.h. in welchem Maß für jedes Items die theoretisch schwierigkeitsbestimmenden Merkmale von den Ratern gleich eingeschätzt werden.

Die Frage nach der Einschätzung der Gesamtschwierigkeit des Items wurde sowohl vor als auch nach der Präsentation der Parameter gestellt, so dass zunächst über die *Intraraterreliabilität* betrachtet werden kann, in welchem Grad sich die Einschätzung der Gesamtschwierigkeit nach einer differenzierten, kriteriengeleiteten Analyse verändert. Klassifiziert wurden die Items in drei Stufen (1: leicht, 2: mittel, 3: schwer).

Da es hier um weit weniger Kategorien pro Fall geht – nämlich nur 2, sollte hier der konservative Kappa-Koeffizient verwendet werden.

Die Gleichung zu Cohens Kappa lautet:

$$\kappa = \frac{p_0 - p_c}{1 - p_c}$$

Dabei ist p_0 die gemessene Übereinstimmung der beiden Rater und p_c die Zufallsübereinstimmung. Sollten die Rater in jedem ihren Urteilen exakt übereinstimmen, wäre $\kappa=1$. Sofern sich nur Übereinstimmungen zwischen den beiden Ratern feststellen lassen, die mathematisch dem Erklärungswert des Zufalls entsprechen, nimmt κ einen Wert von 0 an. Es kann auch zu negativen Werten kommen, diese weisen dann auf eine Übereinstimmung hin, die kleiner als die Zufallswahrscheinlichkeit ist. Der Vorzug von κ ist also, dass die Erklärungsmacht des Zufalls im Ergebnis berücksichtigt wird (vgl. Bortz et al. 2000, S. 458f).

Zur Einordnung des Wertes schlagen sodann Landis/Koch (1977) folgende Werte vor:

Tabelle 29 – Bewertung der Kappa-Statistiken nach Landis/Koch (1977)

κ-Statistiken	Stärke der Übereinstimmung
< 0.00	Poor [schwach]
0.00-0.20	Slight [etwas]
0.21-0.40	Fair [ausreichend]
0.41-0.60	Moderate [mittelmäßig]
0.61-0.80	Substantial [beachtlich]
0.81-1.00	Almost Perfect [fast vollkommen]

Obwohl sie selbst zugeben: “Although these divisions are clearly arbitrary, they do provide useful ‘benchmarks’ for the discussion...”¹¹⁴ (Landis/Koch 1977, S. 165).

¹¹⁴ „Obwohl diese Einteilungen offensichtlich willkürlich sind, bieten sie doch einige nützliche Werte für die Diskussion“

Bei Rater A betrug die diesbezügliche Intraraterreliabilität und damit die Urteilsstabilität $\kappa^{115} = .628$.

Tabelle 30 – Intraraterreliabilität Rater A

	Wert	Asymptotischer Standardfehler ^a	Näherungsweise T ^b	Näherungsweise Signifikanz
Maß der Übereinstimmung Kappa	,628	,079	7,236	,000
Anzahl der gültigen Fälle	65			

a. Die Null-Hypothese wird nicht angenommen.

b. Unter Annahme der Null-Hypothese wird der asymptotische Standardfehler verwendet.

Somit kann der Wert von Rater A als beachtlich bewertet werden, was meint, dass die Übereinstimmung unabhängig von zufälligen Übereinstimmungen, die bei Cohen wie dargestellt berücksichtigt werden, eine gute Übereinstimmung zeigt. Daher kann davon ausgegangen werden, dass Rater A nicht entscheidend in seinem Urteil abweicht, unabhängig davon, ob die Frage nach der Gesamtschwierigkeit vor oder nach der Parameterschätzung gestellt wird.

Bei Rater B hingegen ist $\kappa = .526$ und damit eher als mittelmäßig zu bewerten, wenngleich der Wert zumindest in der oberen Hälfte des Bereich liegt.

Tabelle 31 – Intraraterreliabilität Rater B

	Wert	Asymptotischer Standardfehler ^a	Näherungsweise T ^b	Näherungsweise Signifikanz
Maß der Übereinstimmung Kappa	,526	,089	5,631	,000
Anzahl der gültigen Fälle	65			

a. Die Null-Hypothese wird nicht angenommen.

b. Unter Annahme der Null-Hypothese wird der asymptotische Standardfehler verwendet.

Dies lässt den Schluss auf eine leichte Beeinflussbarkeit von Rater B durch die Teilfragen zu. Dies ist nicht als problematisch zu sehen, sollte jedoch beim späteren Vergleich von Punktwerten als Maß der Schwierigkeit und der intuitiven Einschätzung berücksichtigt werden.

Eine Analyse der Mittelwerte (Schwierigkeitseinschätzung vorher und Schwierigkeitseinschätzung nachher) ergab, dass Differenzen der beiden Mittelwerte der Einschätzungen nur .02 beträgt (eine 0 würde keinerlei Abweichungen bedeuten). Die Ergebnisse der Bildung der Differenzen der Schwierigkeitseinschätzungen, d.h. der Subtraktion des Werts von Einstufung vorher vom Wert der Einstufung nachher, sind in Tabelle 32 abgetragen. Sie zeigen, dass nur 2 Mal Abweichungen (entspricht 3% der Fälle) einen Wert von -2 oder 2 annehmen, d.h. eine Kategorie übersprungen haben. Damit sind die Einschätzungen insgesamt als konsistent, nicht zufällig und verhältnismäßig stabil zu bewerten.

¹¹⁵ κ =Cohens Kappa

Tabelle 32 – Abweichungen der Einschätzungskategorien Rater B

	Häufigkeit	Prozent
-2 (große Abweichung nach oben)	1	1,5
-1 (kleine Abweichung nach oben)	8	12,3
0 (keine Abweichung)	46	70,8
1 (kleine Abweichung nach unten)	9	13,8
2 (große Abweichung nach unten)	1	1,5
Gesamt	65	100,0

Auf Grund des Vorgenannten kann insgesamt davon ausgegangen werden, dass es nicht zu groben Verzerrungen kommt, wenn im Folgenden lediglich die nachgeordnete Einschätzung betrachtet wird. In Summa scheint auch eine Beeinflussung der Einschätzung der Gesamtschwierigkeit durch einer differenzierten, kriteriengeleiteten Analyse nicht problematisch, sondern naheliegend, was ebenfalls für die Verwendung der nachfolgenden Einschätzung spricht. Bei weiteren diesbezüglichen Ratings kann also davon ausgegangen werden, dass eine Platzierung der Einschätzung nach der Analyse sinnvoll ist.

Jedes einzelne schwierigkeitsbestimmende Merkmal bzw. die dazugehörigen Antworten wurden je nach Schwierigkeitsstufe aufsteigend bepunktet, so dass schließlich ein Punktwert für jedes Items errechnet werden kann (siehe Tabelle 33). Dabei ergibt sich ein Spektrum erreichbarer Punkte von min. 8 max. 22 Punkten, wobei eine hohe Punktzahl eine höhere Schwierigkeit bedeutet.

Eine Reliabilitätsberechnung bezüglich des Zusammenhangs der Punkte der beiden Rater ergab nur nicht-signifikante Werte, die somit nicht weiter interpretiert werden können.

Wie jedoch an Hand von Tabelle 33 erkennbar ist, stimmen die Punktzahlen der Rater bei 7 Items vollständig überein (Items 2, 21, 23, 37, 45, 49 und 67¹¹⁶). Die größte und unikale Differenz von 8 Punkten lag bei Item 62 vor (weitere diesbezügliche Analysen siehe unten). Die mittlere Differenz beträgt 2 und kann somit als gering angesehen werden, so dass auf dieser basalen Ebene des Anscheins festgehalten werden kann, dass die mit Hilfe des Ratings erzielten Punkte eine gewisse Stabilität zwischen den Ratern aufweisen.

Tabelle 33 – Punkteverteilung der Items nach dem Rating

ID	Summe Rater A	Summe Rater B
1	16	15
2	8	8
3	14	12
4	14	12
5	15	11
6	8	11

¹¹⁶ Angegeben ist hier und im Folgenden immer die Item ID

7	20	19
8	14	13
9	21	17
10	15	12
11	22	19
12	16	11
13	15	17
14	17	13
15	11	13
16	12	11
17	13	14
18	14	10
19	14	11
20	17	16
21	16	16
23	16	16
24	13	15
25	12	10
27	18	12
28	15	12
29	15	12
32	12	13
33	12	10
34	18	13
35	20	15
36	17	13
37	10	10
38	17	12
39	12	11
40	16	14
41	19	18
42	20	19
43	18	12
45	14	14
46	17	14
47	20	14
48	14	12
49	12	12
50	11	10
51	13	12
52	15	13
54	19	17
55	14	10
56	17	15
57	16	13
58	18	15

59	14	11
60	15	12
61	16	13
62	22	14
63	17	15
64	18	17
66	16	12
67	13	13
68	12	11
69	15	10
70	16	12
72	13	12
73	14	15

Um diese Behauptung zu belegen, wurde überprüft in wie weit die Rater in ihren Punktwerten korrelieren.

Folgende Richtwerte werden als Hilfe für die Interpretation eines Korrelationskoeffizienten nach Pearson empfohlen – analog zu dem bezüglich des Kappa-Werts geschriebenen, sind auch dies keine klaren Grenzwerte, sondern vielmehr Orientierungen (vgl. Brosius 1998, S. 503).

Tabelle 34 – Bewertung des Korrelationskoeffizienten nach Brosius (1998)

Betrag des Korrelationskoeffizienten	Möglich Interpretation
0	Keine Korrelation
über 0 bis .2	Sehr schwache Korrelation
.21 bis .4	Schwache Korrelation
.41 bis .6	Mittlere Korrelation
.61 bis .8	Starke Korrelation
.81 bis unter 1	Sehr starke Korrelation
1	Perfekte Korrelation

Die Punkte der Rater A und B korrelieren signifikant bei .69, was als starke Korrelation interpretiert werden kann.

Tabelle 35 – Korrelation der Punktwerte der Rater

	Punktwert-A	Punktwert-B
Korrelation nach Pearson	1	,697**
Punktwert-A Signifikanz (2-seitig)		,000
N	65	65
Korrelation nach Pearson	,697**	1
Punktwert-B Signifikanz (2-seitig)	,000	
N	65	65

**. Die Korrelation ist auf dem Niveau von .01 (2-seitig) signifikant.

In diesem Fall wird allerdings nicht angenommen, dass die beiden Einschätzung sich gegenseitig beeinflussen, sondern, dass beide Einschätzungen von einem dritten Faktor – der angenommenen Schwierigkeit der Items ausgedrückt durch die Parameter – abhängen.

Diese Ergebnisse betrafen zunächst nur die globale Schwierigkeitsbestimmung über alle Items und Merkmale hinweg. Im Folgenden werden sowohl einzelne Items als auch einzelne Merkmale genauer beleuchtet.

Besonders interessant ist zunächst die Betrachtung der Übereinstimmung über alle Merkmale pro Item zwischen Rater A und Rater B. Zur Erläuterung: Hier wurde überprüft, wie für jedes der 65 Items die 12 Merkmale geratet wurden und wie die Übereinstimmung dieses Rating aussehen. Ziel der Überprüfung der *Interraterreliabilität* ist die Sicherstellung der Replizierbarkeit des Kategoriensystems mittels der Überprüfung der Koderierabhängigkeit der Schwierigkeitsbestimmenden Merkmale. Auch hier wird aus den gleichen Gründen wie oben genannt (hohe Anzahl an Kategorien) der Holsti-Koeffizient verwendet sowie die prozentuale Übereinstimmung angegeben. Die Ergebnisse sind in Tabelle 36 zu finden.

Tabelle 36 – Raterreliabilitäten Itemebene

Item ID	Prozentuale Übereinstimmung	R
1	75%	0.92
2	100%	1
3	75%	0.92
4	66.7%	0.75
5	91.7%	0.92
6	83.3%	0.92
7	75%	0.75
8	58.3%	0.67
9	66.7%	0.67
10	66.7%	0.83
11	66.7%	0.67
12	50%	0.75
13	83.3%	0.83
14	66.7%	0.83
15	66.7%	0.83
16	83.3%	0.83
17	58.3%	1
18	66.7%	0.83
19	66.7%	0.67
20	75%	0.92
21	75%	0.92
23	66.7%	0.83
24	100%	1
25	83.3%	0.83

27	66.7%	0.83
28	91.7%	0.92
29	58.3%	0.75
32	75%	0.75
33	91.7%	0.92
34	50%	0.67
35	75%	0.75
36	66.7%	0.83
37	91.7%	0.92
38	66.7%	0.67
39	75%	0.83
40	58.3%	0.83
41	91.7%	0.92
42	83.3%	0.83
43	50%	0.67
45	66.7%	0.67
46	58.3%	0.75
47	58.3%	0.75
48	75%	0.75
49	100%	1
50	75%	0.75
51	58.3%	0.75
52	50%	0.83
54	50%	0.67
55	66.7%	0.92
56	41.7%	0.58
57	75%	0.92
58	75%	0.83
59	75%	0.83
60	50%	0.75
61	58.3%	0.92
62	50%	0.50
63	58.3%	0.58
64	75%	0.75
66	58.3%	0.75
67	100%	1
68	91.7%	0.92
69	58.3%	0.75
70	58.3%	0.75
72	66.7%	1
73	83.3%	0.83

Wie zu erkennen ist, sind die Werte insgesamt zufriedenstellend, wenn auch nicht herausragend. Das arithmetische Mittel der prozentuale Übereinstimmung liegt bei 71%, der Median der prozentualen Übereinstimmung bei 67%, der durchschnittlichen Wert R-Wert beträgt .81,

der Median liegt bei .83. Ab.80 wird dieser als akzeptabel gewertet (vgl. Lauf 2001, S. 65f). Daher kann das Kategoriensystem zwar nicht als optimal, aber doch als ausreichend befunden werden, um eine erste Einschätzung der Items vorzunehmen.

Werden in dieser Tabelle die Extreme betrachtet, so fällt positiv auf, dass bei 6 Items (ID: 2, 17, 24, 49, 67, 72) der Holsti-Koeffizient einen Wert von 1 annimmt und bei 4 Items (2, 24, 49, 67) eine 100%ige Übereinstimmung erreicht wird. Den Gegensatz dazu bildet – sowohl was eine der geringsten prozentualen Übereinstimmung (50%) und was den niedrigsten Holsti-Wert betrifft (.50) – Item 62. Dieses Item weist dazu konsistent auch die größte Punktedifferenz auf. Aus diesem Grund und da es in Kapitel 4.1.4 (S. 159) bereits ausführlich vorgestellt wurde, scheint eine detaillierte Betrachtung an dieser Stelle lohnend.

Einigkeit besteht bei beiden Ratern bezüglich der Merkmale

- Anzahl der PowerPoint Folien/Seiten = 2
- Art und Anzahl der Informationsquellen = eine Quelle
- Das Format des Itemstimulus ist... = depiktional
- Anzahl der irrelevanten/konkurrierenden Informationen im präsentierten Material = viele
- Die Antworten auf das Item erfolgen = geschlossen mit drei plausiblen Antwortalternativen
- Notwendigkeit besonderen Weltwissens = nein

Abweichungen ergeben sich in den Punkten:

- Es müssen Inferenzen gebildet werden = A: viele | B: keine
- Moduswechsel zur Beantwortung notwendig = A: ja | B: nein
- Komplexität des Mentalen Modells = A: komplex | B: kein Modell
- Platzierung der relevanten Informationen = A: versteckt | B: offen
- Notwendigkeit besonderen Sprachwissens = A: ja | B: nein

Insgesamt schätzte Rater A das Item als „schwer“ ein – Rater B eher als „mittel“.

Die Einschätzung von Rater B verwundert, da beim Transfer einer depiktionalen Repräsentation in eine deskriptionale Aussage eigentlich von einem Moduswechsel ausgegangen werden kann und dieser Wechsel ein mentales Modell voraussetzt. Bei der Menge der zu verarbeitenden und aufeinander zu beziehenden Informationen, daher auch die Annahme eine große Anzahl von Inferenzen müssen gebildet werden, wurde dieses Modell von Rater A als komplex eingestuft, wohingegen als Konsequenz aus der nicht Notwendigkeit eines Moduswechsels Rater B logisch korrekt auch hier angab, dass kein Modell erforderlich sei, was ein Versehen seitens Rater B

unwahrscheinlich macht, da dieser ebenfalls logisch konsistent angab, dass keine Inferenzen gebildet werden müssen.

Diese in Kernpunkten unterschiedlichen Betrachtungen führen zu dem großen Unterschied bei der Punkteverteilung und erklären die schlechten Reliabilitäts- und Prozentwerte. Weiterführend wäre es daher interessant solche Ratings zukünftig mit Begründung der Rater durchzuführen – in vorliegenden Fall wurde dies aus testökonomischen Gründen (65 Items x 12 Merkmale) nicht durchgeführt, dadurch ist jedoch die Interpretation der vorliegenden Ergebnisse erschwert. Bei folgenden Itemkonstruktionen und entsprechenden Ratings scheint es daher ratsam zumindest exemplarisch oder bei besonders wichtigen Items (Ankeritems, Niveauschwellen) die Ratings zumindest kurz begründen zu lassen bzw. im Anschluss Diskussionen über abweichende Einschätzungen zu führen.

An Hand von Tabelle 37 lässt sich feststellen, dass Item 11 und Item 42 von beiden Ratern gemäß der Punkte¹¹⁷ als besonders schwer eingestuft wurde. Rater A klassifizierte weiterhin auch Item 62 sowohl bezogen auf die Parametereinschätzung als auch auf die Gesamteinschätzung als besonders schwer, genauso wurden gemäß der Punkte Item 9 und Item 47 als schwierig eingestuft. Dies trifft bei Rater B für Item 7, 41 und Item 64 zu. Item Nummer 2 wird von beiden Ratern einhellig in jeder Form der Schwierigkeitseinschätzung (Punkte wie Intuitiv) als besonders leicht klassifiziert. Rater A klassifizierte daneben Item 6 ebenfalls eindeutig als leicht.

Tabelle 37 – Vergleich Extremwerte Rater A und B

	Items Rater A	Items Rater B
Leicht	2, 6, 15, 37 , 50	2, 18, 25, 33, 37
Schwer	9, 11 , 42 , 47, 62	7, 11 , 41, 42 , 64

Die Ratingergebnisse bis zu diesem Punkt lassen sich zusammenfassen und aufschlüsseln nach den oben zur Klassifikation der Items verwendeten kognitiven Anforderungsbereichen und den Präsentationsformaten. Daraus ergeben sich folgende Zusammenstellungen (siehe Tabelle 38 und Tabelle 39) – angegeben ist jeweils der Mittelwert der betreffenden Betrachtung.

Tabelle 38 – Ratingergebnisse und Raterreliabilitäten nach Repräsentationsformaten

Deskriptional	Hybrid	Depiktional
Punkte: 12.62	Punkte: 14.62	Punkte: 14.85
Holsti: .83	Holsti: .79	Holsti: .80
Prozent: 74.12%	Prozent: 61.68%	Prozent: 69.09%

¹¹⁷ bei Rater A auch gemäß der intuitiven Einschätzung

Tabelle 39 – Ratingergebnisse und Raterreliabilitäten nach kognitiven Anforderungsbereichen

identifizieren	Integrieren	generieren
Punkte: 12.25	Punkte: 13.47	Punkte: 16.26
Holsti: .83	Holsti: .83	Holsti: .79
Prozent: 74.95%	Prozent: 65.61%	Prozent: 72.63%

Erkennbar ist, dass es zwischen den Repräsentationsformaten keine sehr großen Unterschiede bezüglich der Punkte gibt, auch die Raterreliabilitäten bewegen sich alle in einem ähnlichen Bereich.

Für die kognitiven Anforderungsbereiche hingegen zeigen die Punkte erwartungskonform eine aufsteigende Tendenz.

Bereits ausgehend von den gemachten Beobachtungen hat es den Anschein, dass es einen Zusammenhang zwischen den kognitiven Anforderungsbereichen und der Schwierigkeitseinschätzung via gibt.

4.2.4. Exkurs III – Analyse der Itemschwierigkeit unter Aspekten der Verständlichkeitsforschung II

Unabhängig von den Ergebnissen des Ratings ist eine empirische Überprüfung der theoretisch schwierigkeitsbestimmenden Parameter angebracht. Dabei wird kritisch reflektiert, ob diese Merkmale sich zukünftig als Apriori Bestimmungsgrößen der Itemschwierigkeit eignen und eingesetzt werden können.

Erkenntnisperspektive, Forschungshypothese und Untersuchungsdesign

Erkenntnisleitend bei den folgenden Analysen ist die Frage:

Ist auf Basis der oben angenommenen schwierigkeitsbestimmenden Faktoren möglich, bereits bei der Konstruktion von Items verschiedene Schwierigkeitsabstufungen berücksichtigen zu können, und insgesamt die Anforderungen über die aus der Theorie abgeleiteten Parameter bestimmen zu können?

Die dahinterstehende *Hypothese* ist, dass die verschiedenen Itemgruppen innerhalb der Kategorien (die im Folgenden vorgestellt werden) signifikant in der theoretisch angenommenen Richtung differieren.

Um dies zu untersuchen wurden alle Items an Hand der oben ausgeführten Kategorien klassifiziert¹¹⁸. Anschließend wurden die mittleren Itemschwierigkeiten je Kategorie ermittelt, eine Varianzanalyse bzw. ein t-Test durchgeführt und insofern das Ergebnis signifikant war, die Passung auf das theoretische Modell betrachtet. Bei nicht signifikanten Ergebnissen wurde eine Anscheinsanalyse durchgeführt, die jedoch keine grundsätzlichen Aussagen erlaubt. Die weiteren Analysen beziehen sich alle auf die CBA Version des Tests.

Einfluss der theoretisch hergeleiteten Parameter auf die Schwierigkeit der Items

Um die mittleren Schwierigkeiten der Items bestimmen zu können wurden alle Items eindeutig einer möglichen Ausprägung des jeweiligen Parameters zugeordnet, dabei wurde offenkundig, dass bezüglich der Länge des einzelnen Items als beeinflussender Faktor der Vergleich problematisch ist, da es 41 Items mit 2 Folien Länge gibt, aber bspw. nur eins mit der Länge von 4 Folien, so dass ein Vergleich nicht sinnvoll erscheint und daher unterbleibt.

Ein ähnliches Problem tritt bei den Merkmalen „Sprachwissen“ und „Weltwissen“ auf – es sind in einer der Ausprägungen jeweils zu wenig Items vorhanden, so dass auch hier ein Vergleich nicht zielführend scheint.

Die anderen Merkmale und ihre jeweiligen Werte sind in der Tabelle angeführt – einige mussten, um genug Items für jede Kategorie vorweisen zu können, in ein dichotomes Betrachtungssystem überführt werden (Inferenzbildung: ja | nein; Bildung eines mentalen Modells: ja | nein; Platzierung der relevanten Informationen: offen | versteckt; Menge der irrelevanten Informationen: wenig | viel).

Tabelle 40 – Mittelwerte der Estimates der approximierten schwierigkeitsbestimmenden Merkmale

Merkmal	Gruppe	Estimate
Art + Anzahl der Quellen	Eine Quelle	-.61
	mehrere Quellen des gleichen Repräsentationsmodus	.03
Quellen	mehrere Quellen verschiedener Repräsentationsmodi	-.41
Moduswechsel	Notwendig	-.56
	Nicht notwendig	-.16
Inferenzbildung	Ja	-1

¹¹⁸ Die Klassifikation wurde von zwei Ratern separat vorgenommen und in Fällen von Urteilsabweichungen konnte eine gemeinsame Lösung diskutierend entschieden werden.

	Nein	-.29
Konstruktion eines Mentalen Modells	Ja	.01
	Nein	-.90
Platzierung der relevanten Informationen	Versteckt	.33
	Offen	-.56
Menge der irrelevanten Informationen	Viel	-.26
	Wenig	-.60
Antwortformat	Offen	-.37
	Geschlossen	-.79

Bezüglich der Ergebnisse ist offenkundig, dass das Merkmal „Art und Anzahl der Quellen“ als nicht eindeutig ausschlaggebend für die Schwierigkeit bezeichnet werden kann. Weder ist ein deutlicher Unterschied zwischen einer und mehrere Quellen (-.6 zu -.4) erkennbar noch scheint Multimodalität innerhalb der Quellen einen theoretisch begründbaren Einfluss zu haben (.03 zu -.4). Dies ist nicht kongruent zur Theorie.

Bei dem Merkmal „Moduswechsel“ zeigt sich eine Tendenz, dass Items, die zur korrekten Lösung einen Moduswechsel erfordern, einfacher sind als solche, die es nicht tun. Auch dieser Befund passt nicht zu den theoretischen Annahmen.

Verhältnismäßig eindeutig ist der Befund bei den Merkmalen „Konstruktion eines mentalen Modells“ und „Inferenzbildung“. Sofern die Konstruktion eines mentalen Modells notwendig scheint, sind die Items durchschnittlich um .91 schwerer. Das ist kongruent zur Theorie, da davon ausgegangen werden kann, dass die mentale Modellierung eines Sachverhalts, wie bereits dargestellt, kognitiv relativ aufwendig ist und größere Kompetenz erfordert als die darunter liegenden Stufen. Auch die Inferenzbildung scheint einen, wenn auch schwächeren, Einfluss auf die Schwierigkeit auszuüben (.71). Auch dies ist kongruent zur Theorie.

Eine ähnlich starke Ausprägung (.89) besitzt das Merkmal „Platzierung der relevanten Informationen“ in der Form, dass offen platzierte Informationen die Beantwortung des Items deutlich vereinfachen. Diese Beobachtung weist eine Passung zur Theorie auf.

Die Menge der irrelevanten Informationen dagegen scheint wiederum eher von untergeordneter Bedeutung zu sein. Wenngleich Items mit weniger irrelevanten Informationen passend zur Theorie auch einen geringere Schwierigkeit aufweisen.

Auch das Antwortformat zeigt erwartungskonforme Ergebnisse: geschlossene Items sind leichter zu beantworten als offene.

Gründe für erwartungswidrige Ergebnisse können einerseits in der Konstruktion der Items liegen, was meint, dass diese Faktoren in den Items nicht in Reinform vorkommen und es so zu

Wechselwirkungen der Faktoren kommen kann, oder andererseits in der inhomogenen Itemzahl pro Gruppe.

Bezüglich der Aussagekraft der obigen Analysen muss jedoch festgehalten werden, dass T-Tests und Varianzanalysen außer bei dem Merkmal „Konstruktion eines Mentalen Modells“ nur nicht signifikante Ergebnisse ergaben, was die obige These stützt, dass auf Grund der Inhomogenität der Items die Mittelwertsunterschiede zufällig zustande gekommen sind. Einzig das Merkmal „*Konstruktion eines Mentalen Modells*“ weist bei Varianzgleichheit einen signifikanten T-Wert auf ($T = -3.5$, $p = < .001$, $df = 63$). Bezüglich der Effektstärke (cohens $d = .88$) liegt gemäß den Konventionen von Cohen (vgl. Cohen 1992, S. 157) ein großer Effekt vor. Dies ist kongruent zur dargelegten Theorie und unterstützt die Relevanz dieses Faktors als schwierigkeitsbestimmender Parameter. Damit kann zumindest festgehalten werden, dass Items, die vom Probanden eine mentale Modellierung des Sachverhalts fordern, signifikant schwieriger sind als solche, die es nicht erfordern.

Insgesamt problematisch bei dieser Betrachtung ist wie oben angedeutet die Tatsache, dass die Merkmale in den Items nicht kontrolliert sind, das bedeutet, dass mehrere Faktoren in einem Item zusammen kommen. Damit lässt sich eine Aussage wie „Sofern die relevanten Informationen versteckt innerhalb des Items platziert werden, kann eine um .89 geringere Lösungswahrscheinlichkeit angenommen werden“ in dieser Form nicht postulieren. Um eine größere Sicherheit, Präzision und Tragfähigkeit solcher und ähnlicher Aussagen zu erreichen, müssten Items geschaffen werden, die sich nur in dem zu untersuchenden Merkmal unterscheiden. Dies sollte auch als Forschungsdesiderat erkannt und bei bzw. vor der Konstruktion weiterer Tests hinreichend untersucht und eruiert werden.

Nach den differenzierten jedoch eher inhaltsbasierten Analysen der Testitems und der Hinterfragung der Konstruktion bleiben noch die Einsatzmöglichkeiten und Adäquanz des Test – also seine anwendungsbezogene Seite zu untersuchen. Ferner bleibt die Frage nach weiterer Anschlussforschung bzw. die sich auch den Ergebnissen abgeleiteten weiteren Fragestellungen.

4.2.5. Exkurs IV – Vergleich der Verfahren zur Schwierigkeitsbestimmung

Im Anschluss an diese Betrachtungen bleibt noch die kurze Frage offen, in wie weit die beiden Verfahren (Ratingpunkte und CBA Kalibrierung) zur Schwierigkeitsbestimmung der Items insgesamt ähnliche Ergebnisse liefern.

Dazu erfolgt eine Korrelation der im Rating an Hand des Mittelwerts der Punkte beider Rater ermittelten Schwierigkeiten und den in der CBA Kalibrierung ermittelten Schwierigkeiten:

Tabelle 41 – Korrelationen der Schwierigkeiten Rating + CBA Kalibrierung

		Schwierigkeit (CBA Kalibrierung)	Schwierigkeit (Rating)
Schwierigkeit (CBA Kalibrierung)	Korrelation nach Pearson	1	,308*
	Signifikanz (2-seitig)		,013
	N	65	65
Schwierigkeit (Rating)	Korrelation nach Pearson	,308*	1
	Signifikanz (2-seitig)	,013	
	N	65	65

*, Die Korrelation ist auf dem Niveau von .05 (2-seitig) signifikant.

Zwar ist die Korrelation signifikant, jedoch allenfalls als schwach bis mittel zu beurteilen.

Somit muss konstatiert werden, dass an Hand dieser Ergebnisse und der Ergebnisse aus Kapitel 0 das vorliegende Ratingverfahren und die ermittelten Parameter als a priori Schwierigkeitsbestimmung von Items nicht geeignet scheint und nur Annäherungen bringen.

4.2.6. Niveaustufenmodellierung

Rekapitulation: Skalierung des Tests

Die Skalierung der MaK-adapt/Lesen Daten erfolgte wie bereits ausgeführt auf Grundlage der Item-Response-Theorie, hierbei werden die Antwortwahrscheinlichkeiten der einzelnen Aufgaben als eine Funktion der zugrundeliegenden Fähigkeit betrachtet. Das in MaK-adapt/Lesen verwendete IRT-Modell ist ein generalisiertes Rasch-Modell, welches in der Analysesoftware ConQuest implementiert ist. Die Auswertung auf Basis des Rasch-Modells hat eine Reihe untersuchungs- und auswertungstechnischer Vorteile.

Im Rahmen einer Rasch-Skalierung der Leistungsdaten ist es möglich, für Auszubildende, die unterschiedliche Items bearbeitet haben, Schätzwerte auf einer gemeinsamen Skala zu ermitteln. Ein weiterer Vorteil der Skalierung der konstruierten Leistungstests mittels eines Rasch-Modells ist die Möglichkeit, Aufgabenschwierigkeiten und Personenfähigkeit auf einer gemeinsamen Skala abzubilden. Bei einer klassischen Testwertbildung hingegen lässt sich zwischen der Kompetenz eines Auszubildenden (wie viel Prozent aller Aufgaben er lösen konnte) und der Schwierigkeit einer Aufgabe (von wie vielen Prozent der Auszubildenden wurde sie gelöst) keine Beziehung herstellen. Die Raschskalierung dagegen bildet beide Größen auf eine gemeinsame Skala ab, wobei der Bezug zwischen der Kompetenz der Auszubildenden und der Schwierigkeit der Items in Form von Lösungswahrscheinlichkeiten hergestellt wird. Die Schwierigkeit eines Items ist im Rasch-Modell definiert als Punkt auf der Kompetenzskala, an dem die Wahrscheinlichkeit der Probanden mit der Kompetenz θ , das Item zu lösen, bei 50% liegt.

Anhand dieses dem Rasch-Modell zugrunde gelegten Zusammenhanges zwischen Personenfähigkeit und Lösungswahrscheinlichkeit lassen sich jedoch auch Punkte auf der Kompetenzskala bestimmen, an denen die Lösungswahrscheinlichkeit für ein Item einen anderen Wert als 50% annimmt. „Bei Beschreibung von Kompetenzen interessiert, ob eine Population die Anforderungen bestimmter Aufgaben mit einer hinreichenden Sicherheit bewältigen kann – eine Lösungswahrscheinlichkeit von 50% erscheint hierfür relativ niedrig. Daher werden oft höhere Werte gewählt, um einzelne Aufgaben auf der Kompetenzskala zu verorten.“ (Hartig 2007, S. 85)

In MaK-adapt/Lesen sind die Skalen an Hand einer 65%-Schwelle, wie sie bspw. in TIMSS oder DESI verwendet wurden, charakterisiert.

Aus den vorliegenden Kompetenzdaten resultieren quantitative Ergebnisse auf kontinuierlichen Skalen. Diese Zahlenwerte sind gut geeignet, um Zusammenhänge der erfassten Kompetenzen mit anderen Variablen zu untersuchen (bspw. Geschlecht oder dem schulischen Abschluss) oder die Kompetenzen verschiedener Gruppen (bspw. verschiedener Berufe) zu vergleichen. Neben derartigen quantitativen Zusammenhangsanalysen ist es aber von Interesse, über welche spezifischen Kompetenzen Auszubildende auf einem bestimmten Niveau verfügen bzw. welche Leistungsanforderungen allgemeiner Kompetenzen sie bewältigen können.

Somit besteht Bedarf an einer kriteriumsorientierten Darstellung und Auslegung der quantitativen Leistungswerte. Dabei sollen die numerischen Werte der Kompetenzskala zu konkreten, inhaltsbezogenen Anforderungen in Bezug gesetzt werden. Dies soll mit der Bildung so genannter Kompetenzniveaus erreicht werden.

Da es in der Praxis nicht realisierbar ist, für jeden singulären Punkt einer kontinuierlichen Skala eine Beschreibung der jeweiligen Kompetenz vorzunehmen (Beaton/Allen 1992) wird eine Unterteilung dieser Skala in Abschnitte vorgenommen, welche im Folgenden als Kompetenzniveaus bzw. Niveaustufen bezeichnet werden. Für jede dieser Niveaustufen wird dann eine kriteriumsorientierte Beschreibung der jeweiligen Kompetenzen vorgenommen.

Um für die Kompetenzskala des pilotierten Tests Kompetenzniveaus zu bestimmen, stehen verschiedene Methoden zur Verfügung. Die meisten Vorgehensweisen stimmen darin überein, dass für die Definition dieser Niveaustufen die Bestimmung der Schwellen zwischen den Niveaus entscheidend ist (vgl. Hartig 2007, S. 86). Innerhalb eines Skalenabschnittes, der als ein Kompetenzniveau betrachtet wird, wird dann keine weitere inhaltliche Differenzierung der Kompetenzen mehr vorgenommen.

Als inhaltliche Grundlage für die Bildung und die Beschreibung von Skalenabschnitten werden klassischerweise die Schwierigkeiten der Aufgaben eines Tests und deren fachbezogene Anforderungen herangezogen – diese Vorgehens wird auch bei MaK-adapt/Lesen verwendet und im Folgenden dargestellt. Anhand der Items werden die Kompetenzen der Auszubildenden beschrieben, deren Leistung innerhalb des Niveaus liegt.

Definition von Kompetenzniveaus

Formulierung von globalen Itemmerkmalen

Es wurde bereits vor der Verankerung der Items auf der Kompetenzskala globale Klassifikationen der Testaufgaben hinsichtlich möglicher schwierigkeitsbestimmender Charakteristika vorgenommen. Diese Beschreibungen wurden aus dem Modell der zu erfassenden Kompetenzen abgeleitet und beinhalten Annahmen darüber, welche spezifischen Anforderungen zur Schwierigkeit einer Aufgabe beitragen. Zentral sind an dieser Stelle die sogenannten kognitiven Anforderungsbereiche (identifizieren, integrieren und generieren) und das Präsentationsformat (deskriptional, hybrid oder depiktional). Weitere Merkmale sind:

- der gesuchte Begriff bzw. die Aussage wird in der Frage explizit genannt → es müssen nur identische Wörter identifiziert werden vs. benannter und gesuchter Begriff unterscheiden sich
- Anzahl der irrelevanten (ggf. ablenkenden) Informationen
- Umfang und Informationsdichte des Materials → Grad der Informationsdichte (z. B. einfache versus verschachtelte Tabellen, Komplexitätsgrad von Grafiken, Anzahl an Propositionen in einem Text)
- Anzahl der Informationen, die aufeinander zu beziehen sind → je mehr, desto schwieriger
- Informationsquelle → Daten innerhalb eines Dokuments, Verbindung von Text- und Bild, mehrere Repräsentationen verschiedenen Formats
- Stärke der Verbindung zwischen den Informationen → eindeutig zuordenbar vs. Inferenzen erforderlich
- Platzierung im Dokument
- Komplexität des Mentalen Modells → Anzahl an Informationen, die in das Mentale Modell integriert werden müssen bzw. Anzahl an Propositionen die gebildet werden müssen

Aus diesen anforderungsrelevanten Aufgabenmerkmalen sollte sich also zum Teil im vornherein ableiten lassen, welche Aufgaben leichter oder schwerer sein sollten und worauf diese Schwierigkeitsunterschiede zurückzuführen sind. Diese Aufgabenklassifikation liegt als globale Einschätzungen des ganzen Items vor.

Die Aufgabenklassifikation erleichtert es, die jeweilige zu erfassende Kompetenz auf einer verallgemeinerten Ebene, unabhängig von den konkreten Testaufgaben zu beschreiben. Ein wichtiger Ausgangspunkt für die Identifikation und Beschreibung relevanter Aufgabenmerkmale

sind Hypothesen über die Anforderungen, die beim Bearbeiten und Lösen der Aufgaben bewältigt werden müssen. Aufgabenmerkmale können sich dabei auf unterschiedliche theoretisch angenommene Prozesse beim Lösen, aber auch auf technische Oberflächencharakteristika der Aufgaben beziehen (vgl. Hartig 2007, S. 88).

Die Itemmerkmale des MaK-adapt/Lesen Test sind derart definiert, dass jede Testaufgabe sinnvoll hinsichtlich jedes Merkmals einzustufen ist und sich so eine 9 Felder Matrix ergibt in die jedes Item eingeordnet werden kann. Erste Analysen zeigten jedoch, dass es keine schwierigkeitsrelevante Unterscheidung zwischen den Präsentationsformaten der Items gibt. Da weitere Analysen in diesem Kontext also nicht zielführend sind, wird im Folgenden nur noch die Betrachtung der kognitiven Anforderungsbereiche dargestellt.

Beschreibung der Kompetenzniveaus

Bei der empirischen Analyse der Aufgabenmerkmale und -schwierigkeiten zeigte sich, dass singuläre Merkmale wider Erwarten keinen deutlichen Einfluss auf die Aufgabenschwierigkeiten aufzuweisen schienen, was sich nicht einfach durch die Auswahl der Merkmale, sondern vor allem durch starke Abhängigkeiten (Kollinearitäten) zwischen den Merkmalen erklären lässt. Bei der Itemkonstruktion wurden die schwierigkeitsbestimmenden Merkmale vielfach in Kombination miteinander realisiert, so dass Items häufig gleichzeitig hinsichtlich mehrerer Merkmale als leichter oder schwerer einzustufen waren.

So wurden die Niveaustufen primär über eine Inspektion einer ordinalen Reihung der Items gemäß der Itemschwierigkeit durchgeführt. Beginnend vom leichtesten Item wurden dann inhaltliche Sprünge gesucht (z.B. gemäß Facetten des theoretischen Kompetenzmodells). An solchen sind dann die Grenzen zwischen Kompetenzniveaus zu setzen. Ziel war die Bildung von drei bis vier Kompetenzstufen auf der Basis der Kalibrierungsdaten. Als Response-Probability wurde $p = .65$ festgelegt.

Es ergaben sich für die Leseitems vier Kompetenzstufen, die eine unterschiedliche Anzahl von Items umfassten.

Auf der *ersten Stufe* befinden sich 10 Items¹¹⁹. Diese zeichnen sich dadurch aus, dass sie zum Großteil (70%) zum kognitiven Anforderungsbereich „identifizieren“ gehören, die beiden zum Bereich „generieren“ gehörenden Items weisen besonders leicht auszuschließende Distraktoren auf, so dass diese wahrscheinlich maßgeblich zur Einordnung auf dieser Stufe (d.h. zu dieser hohen Lösungswahrscheinlichkeit) geführt haben. Zum Teil sind die Begriffe in den Antworten und im Stimulus identisch und werden noch an anderer Stelle wiederholt, so dass die Stärke der aufeinander zu beziehenden Informationen als hoch charakterisiert werden kann. Die Anzahl der irrelevanten Informationen ist bei allen Items eher gering. Insgesamt sind diese Items bezüglich Umfang und Informationsdichte des Materials eher kurz und bezüglich der Informationsquelle kann festgehalten werden, dass sich die Daten zu meist innerhalb eines Dokuments befinden. Weiterhin muss zu meist kein Mentales Modell zur Beantwortung der Frage gebildet

¹¹⁹ L0489211, L0569311, L0029111, L0179221, L0579131, L0599311, L0689311, L0609311, L0069111, L0249321

werden, in den zwei Fällen, in denen es gebildet werden soll (L0579131, L0249321), ist es sehr einfach und umfasst nur wenige zu integrierende Informationen. Jedoch kann bei diesen Items davon ausgegangen werden, dass wahrscheinlich der Großteil der Probanden kein Modell bildet und stattdessen die Entscheidung für eine Antwort vor allem auf Basis der Distraktoren getroffen wird. Diese Stufe weist einen Mittelwert von -2,413 und einen Range von 0,903 auf.

Zur *zweiten Stufe*¹²⁰ erfolgt der größte Sprung der Lösungswahrscheinlichkeiten über alle Items (0,66). Hier befinden sich insgesamt 27 Items, davon gehören 11 zum Anforderungsbereich „identifizieren“ und jeweils 8 zum Anforderungsbereich „integrieren“ und „generieren“. Somit sind 40% der Items dem einfachsten kognitiven Anforderungsbereich zuzuordnen. Dass nur noch identische Phrasen identifiziert werden müssen, kommt nicht mehr vor, d.h. der benannte und gesuchte Begriff unterscheiden sich darin, dass bspw. die Entsprechung eines Begriffes in einer depiktionalen oder deskriptionalen Repräsentation erkannt werden muss. Auch die Anzahl der irrelevanten Informationen wurde stark erhöht, so dass die Items insgesamt einen höheren Umfang und eine höhere Informationsdichte aufweisen. Die Anzahl der aufeinander zu beziehenden Informationen ist höher und wesentlich häufiger ist ein Moduswechsel bzw. die Verbindung von Informationen aus deskriptionalem und depiktionalen Stimulus notwendig. Die Stärke der Verbindung kann immer noch bei allen Items als hoch bezeichnet werden, so dass durchgehend nur wenige Inferenzen notwendig sind. Wenn ein Mentales Modell gebildet werden muss, ist es immer noch auf sehr basalem Niveau und es müssen nur wenige (1-3) Informationen integriert werden. Diese Stufe weist einen Mittelwert von -0,792 und einen Range von 1,207 auf.

Der Sprung zur *dritten Stufe*¹²¹ weniger ausgeprägt (0,125), jedoch zeigt sich bei der Betrachtung der kognitiven Anforderungsbereiche dieser Stufe ein deutlich anderes Bild: nur 3 Items gehören zum Bereich „identifizieren“, 7 zum Bereich „integrieren“ und 17 (d.h. 62%) zum Bereich „generieren“. Auch bei den „identifizieren“ Items müssen nicht nur explizit gleiche oder synonyme Begriffe gefunden werden, sondern aus einem Pool von Informationen verschiedene Daten gegeneinander abgewogen werden, um die im Bezug zur Frage relevante Information korrekt zu identifizieren und zu benennen. Die Anzahl der irrelevanten Informationen kann bei allen Items als hoch klassifiziert werden wie auch insgesamt der Umfang und die Informationsdichte der Items (z.T. 3 Seiten und mehr). Zumeist muss zur Beantwortung der Frage mehr als eine Information und mehr als eine Informationsquelle, häufig in verschiedenen Repräsentationsformaten herangezogen werden. Die Stärke der Verbindung ist hier als eher gering zu beschreiben und die Probanden müssen abstrahieren und Inferenzen bilden. Insgesamt muss hier immer in einem gewissen Maße ein Mentales Modell gebildet werden, welches zum Großteil als komplex bis sehr komplex charakterisiert werden kann, d.h. es muss eine größere Anzahl

¹²⁰ L0349221, L0049221, L0339321, L0089311, L0219231, L0209211, L0109321, L0459311, L0379121, L0739231, L0149221, L0619321, L0729211, L0399121, L0139231, L0299121, L0189221, L0119131, L0699311, L0479131, L0329111, L0279131, L0319231, L0259111, L0649131, L0669121, L0309111

¹²¹ L0059131, L0419231, L0169121, L0429131, L0519121, L0499211, L0269221, L0709131, L0559131, L0589231, L0439221, L0239331, L0679321, L0199221, L0039211, L0359331, L0389121, L0449231, L0469221, L0019221, L0549321, L0289131, L0229131, L0409131, L0539211, L0079231, L0099231

an Propositionen in das Modell integriert werden. Diese Stufe weist einen Mittelwert von 0,257 und einen Range von 0,941 auf.

Stufe 4 weist wiederum einen größeren Sprung zwischen den Items auf (0,283). In dieser Stufe befinden sich 7 Items¹²². Hier bietet der kognitive Anforderungsbereich kaum mehr eine Erklärungsmacht. Hier finden sich Items aus allen drei Anforderungsbereichen. Item L0159221 gehört zum Anforderungsbereich „identifizieren“ und die gesuchte Antwort steht nahezu wörtlich im Text, jedoch ist hier besonders genaues Lesen erforderlich, welches auch ein höheres Maß an Motivation voraussetzt. Bei diesem Item mit offenem Antwortformat gaben viele Probanden nur die Überschrift des betreffenden Abschnitts an, ohne die Antwort korrekt zu differenzieren. Die Items des Bereichs „integrieren“ (L0129321, L0369221, L0529221) weisen alle überaus viele irrelevante Informationen auf und erfordern eine gewisse Problemlösekompetenz im Sinne eines genauen Lesens, da hier die Distraktoren eine hohe Ähnlichkeit aufweisen. Auch die Items L0719121, L0629331 und L0639321 aus dem Bereich „generieren“ erfordern genaues Lesen, um das Material richtig deuten zu können. Bei Item L0639321 kommt weiterhin als Faktor Graphikstrukturwissen hinzu, um alle in der Graphik präsentierten Elemente richtig deuten zu können, aber auch hier führt bereits eine genaue Betrachtung der y-Achse auf die richtige Lösung. Zusammenfassend kann also bei Stufe 4 eine besondere Anforderung an die Genauigkeit mit der das Material rezipiert wird als einflussreicher schwierigkeitsbestimmender Faktor festgehalten werden. Diese Stufe weist einen Mittelwert von 1,407 und einen Range von 0,676 auf.

¹²² L0129321, L0159221, L0719121, L0369221, L0629331, L0529221, L0639321

5. Praxistransfer, Rekapitulation & Ausblick

5.1. Praktischer Einsatzbereich des Tests und Evaluation

Um die anwendungsbezogene Seite des Tests untersuchen zu können, werden einerseits die gemachten Erfahrungen beim Einsatz des Tests in den verschiedenen Erhebungsphasen (an Hand von Erhebungsprotokollen der drei Standorte Darmstadt, Göttingen und Jena, sowie Protokollen von Auswertungsgesprächen) und die in der Pilotierung gestellten Evaluationsitems herangezogen.

Bisher wurde der adaptive Test erfolgreich in den verschiedenen ASCOT-Erhebungen eingesetzt. Bei diesen wird u.a. auch untersucht, in wie weit ein Zusammenhang zwischen beruflichen Kompetenzen und allgemeinen Kompetenzen wie bspw. der Lesefähigkeit besteht. Die bisherigen Erfahrungen im Umgang mit dem Test sind durchweg als positiv zu bewerten. Zwei Aspekte erschweren allerdings zum Teil den Einsatz:

Es ist in einigen beruflichen Bereichen und Schulen überraschenderweise noch der Fall, dass es keine Computerarbeitsplätze in Klassenstärke gibt, so dass entweder für die Testung die Klassen gesplittet werden müssen und es dadurch zu Verzögerungen kommt oder die computerbasierten, adaptiven Tests gänzlich nicht durchführbar sind. Insgesamt ist die technische Kompatibilität des Testprogramms als hoch zu bewerten. Die wenigen Probleme die auftraten (langsame Internetverbindung, kein uneingeschränkter Zugriff auf das Internet,...) konnten alle durch gute Planung und/oder Vorbereitung zufriedenstellend gelöst werden, so dass die eigentlichen Testungen (Kalibrierung und Pilotierung in MaK-adapt) reibungslos verliefen.

Ein anderes Problem scheint die Heterogenität der Schüler in den getesteten Berufen hinsichtlich des Lesetempos zu sein. Es wurden Testungen mit Industriekaufleuten protokolliert, welche 33 Kompetenzitem innerhalb von 20-30 Minuten gut beantworten konnten – hingegen war einige Klassen MFAs oder Altenpfleger innerhalb der gleichen Zeit erst die Bearbeitung von ca. 15-20 Items gelungen. Neben der Lesegeschwindigkeit können natürlich auch andere Faktoren wie die Lesegenauigkeit, Sprachkompetenz oder Motivation eine Rolle spielen. Dies und der Fakt, dass der Test im Grunde nicht als Speedtest konzipiert wurde, macht es schwierig eine Bearbeitungszeit zu empfehlen.

Insbesondere auch für diese Problematik eignet sich die entwickelte adaptive Version. Hier kommt wiederum der Vorteil dieser Art zu Testen zum Tragen, dass sich bereits mit weniger Items höhere Reliabilitätswerte erzielen lassen. Graphisch dargestellt gestaltet sich die (simulierte) Reliabilität der Leseitems im adaptiven Test in etwa wie folgt:

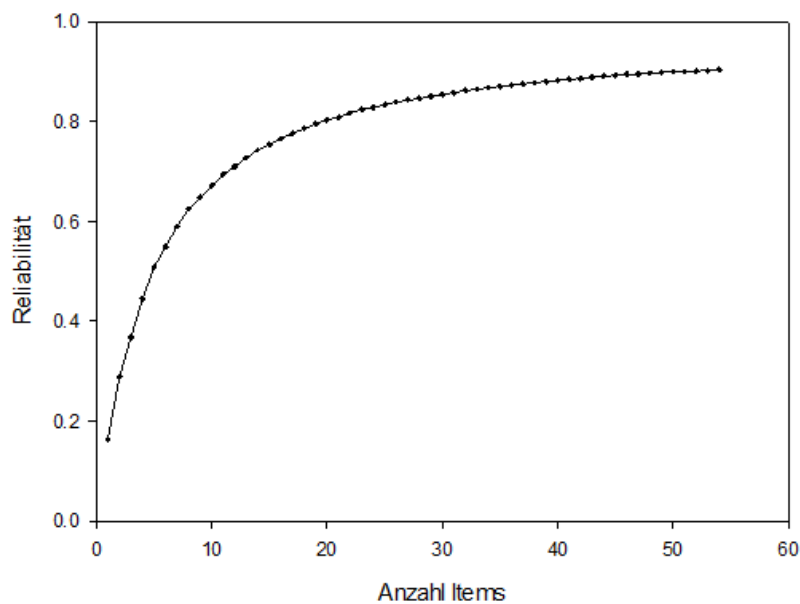


Abbildung 43 – Reliabilitätsentwicklung in Abhängigkeit der Itemanzahl (Simulation) (Frey 2013)

Somit können theoretisch bereits mit 20 Items zufriedenstellende Reliabilitätswerte erreicht werden.

Das vorgestellte Testverfahren zur Ermittlung der Lesekompetenz kann als Gruppentestverfahren eingesetzt werden. Der Test dauert erfahrungsgemäß durchschnittlich eine Schulstunde (à 45 Minuten). Die Auswertung der Schülerantworten kann beim CBA in der Regel computerisiert erfolgen, so dass nur wenige Minuten für die Ermittlung der Leistungsstand der einzelnen Schüler bzw. der ganzen Klasse erforderlich sind.

Bezüglich des Tests insgesamt wurden in der Pilotierungsstudie (Kapitel 0) diverse Evaluationsfragen gestellt, wobei diese vor allem auf die Motivation der Probanden, Umgang mit dem Testmaterial, bevorzugte Testart und wahrgenommene Relevanz der Inhalte (nur PP) abzielten.

Die Motivation wurde über eine 10 stufige Skala (1=wenig; 10=sehr) erfasst. Dabei wurde differenziert zwischen dem vorliegenden Test und einem benotetem Test dessen Leistungen in Bezug auf die Schulnote eines Fachs (bspw. Lesen – Deutsch) berücksichtigt werden. Im Folgenden werden die Ergebnisse sowohl zwischen dem „unbenoteten“ und dem „benoteten“ Test, als auch zwischen CBA und PP Version verglichen.

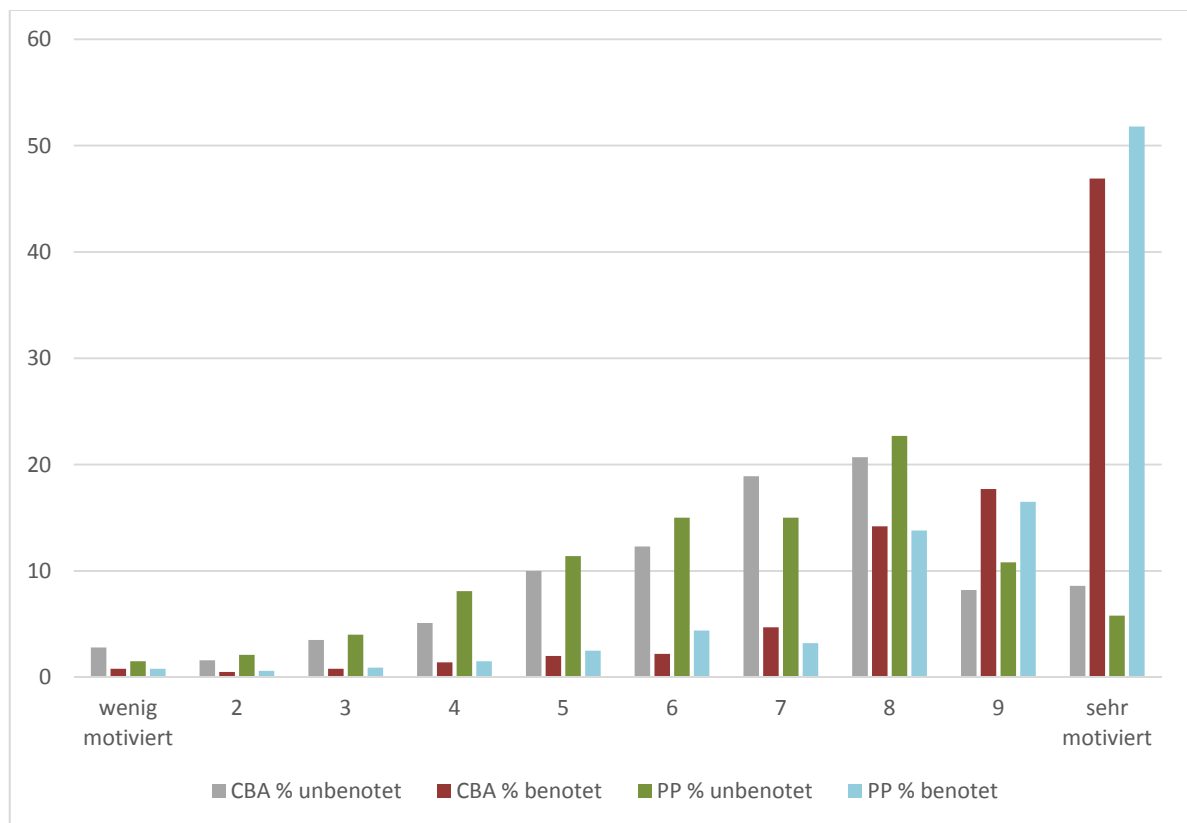


Abbildung 44 - Motivation der Probanden (Selbstauskunft)

Wie Abbildung 44 zu entnehmen ist, ordnen bei einem unbenoteten, computerbasierten Test die meisten Teilnehmer (20,7%)¹²³ ihre Motivation bei 8 ein, bei einem benoteten Test hingegen bei 10 (46,9%)¹²⁴. Insgesamt ist durch einen Vergleich der Mittelwerte (**ub**: 6,7 sd: 2,1; **b**: 8,5 sd: 1,7) zu erkennen, dass bei diesem Test eine Berücksichtigung der Leistung im schulischen Kontext zu einer Steigerung der Motivation geführt hätte.

Bei der PP Version lässt sich eine sehr ähnliche Tendenz erkennen. Ohne Benotung liegt der meist angegebene Wert der Motivation bei 8 (22,7%)¹²⁵ mit Benotung bei 10 (51,8%)¹²⁶. Auch hier lässt ein Vergleich der Mittelwerte (**ub**: 6,6 sd: 2; **b**: 8,9 sd: 1,7) den Schluss zu, dass bei diesem Test eine Benotung zu einer Steigerung der Motivation geführt hätte.

Vergleicht man die Testformate bezüglich ihrer Motivation zunächst bezogen auf einen unbenoteten Test. Fällt auf, dass sie sich nur wenig unterscheiden. Legt man eine Zäsur bei den Werten 5 und 6 an, so ist zu erkennen, dass bei beiden Tests 69% der Personen einen Wert 6 angegeben haben¹²⁷. Dies kann insgesamt als zufriedenstellend bewertet werden, wenngleich sich durch die bloße Anwesenheit der Testleiter eine soziale Erwünschtheit der Antworten nicht

¹²³ 8,3% gaben keine Antwort

¹²⁴ 8,8% gaben keine Antwort

¹²⁵ 3,6% gaben keine Antwort

¹²⁶ 4% gaben keine Antwort

¹²⁷ Es gaben beim PP-Test 27% und beim CBA-Test 23% einen Wert unter 6 an. Die unterschiedlichen Werte bzw. die Differenzen zu 69% ergeben sich aus unterschiedlichen „keine Antwort“-Anteilen.

ausschließen lässt und eine höhere Motivation in Gänze wünschenswert ist. Korreliert man die PP und die CBA Werte jedoch ergeben sich jedoch keine signifikanten Werte.

Ein vergleichbares Bild zeichnet sich auch bei der Berücksichtigung einer Benotung zwischen beiden Testformaten ab. Hierbei geben – bei gleicher Zäsur wie oben – beim PP nur 6,3% und beim CBA 5,5% an wenig motiviert zu sein, hingegen schätzen sich beim PP 90% und beim CBA 85,7% als motiviert ein. Der Einfluss der Berücksichtigung der erreichten Leistung für eine Fachnote auf die Motivation, d.h. die Anstrengungsbereitschaft bei einem Testszenario, scheint also insgesamt unabhängig vom Testformat hoch zu sein. Korreliert man die PP und die CBA Werte jedoch ergeben sich allerdings auch hier keine signifikanten Werte.

Des Weiteren wurde der Umgang mit dem Testmaterial mit folgenden Fragen erfasst, wobei die Antwortmöglichkeiten als quasi-metrische Variable in Form einer vierstufigen Likert-Skala (von 1 „trifft gar nicht zu“ über 2 „trifft etwas zu“ und 3 „trifft ziemlich zu“ bis hin zu 4 „trifft vollkommen zu“) präsentiert wurden.

Die einzelnen Items waren:

- Ich wusste während des Tests jederzeit, was ich zu tun hatte.
- Die Instruktionen waren klar und verständlich.
- Ich war mir während des Tests häufig unsicher, auf welche Weise ich meine *Antwort* angeben soll.
- Der Umgang mit dem Testmaterial war sehr einfach.

Auf Grund der Begrenzung der Testzeit wurde ca. 10% der getesteten Personen diese Fragen nicht mehr vorgelegt, da auch beim adaptiven Test die Kompetenzitems Vorrang vor Evaluationsitems hatten. Fasst man diese vier Variablen zu einer Skala „Das Testmaterial war aus methodischer Sicht einfach zu bearbeiten“ zusammen, so ergibt sich ein α Wert von .74 (PP: .72), so dass diese als gemeinsame Skala interpretiert werden können. Es ergibt sich ein Mittelwert von 2.96¹²⁸ (PP: 3.1¹²⁹), was gerundet den Wert 3 „trifft ziemlich zu“ ergibt und damit die Interpretation zulässt, dass insgesamt in der Stichprobe keine Probleme bei der Bedienung der Testumgebung und bei der Bearbeitung des Tests gab – mit einer minimalen Tendenz, dass der PP bezüglich der Bearbeitung als leichter empfunden wurde –, so dass in dieser Hinsicht nichts gegen den weiteren Einsatz der Items und des Programms spricht.

Ferner wurde danach gefragt, welche Testart (PC oder Papier) bevorzugt würde – dabei gaben 64,7% der computerisiert Befragten an, dass sie die CBA Version bevorzugen würden. Interessanterweise gaben dagegen beim PP-Test 56,4% an, dass sie die Papierversion bevorzugen würden. Woher dieser Effekt kommt ist noch ungeklärt, Vermutungen gehen von einer Beeinflussung durch die vorliegende Testart aus.

¹²⁸ sd: .6

¹²⁹ sd: .61

Abschließend wurden in der PP Version die Probanden bezüglich der Relevanz der Inhalte sowohl für die Berufsschule als auch für die Berufspraxis im Ausbildungsbetrieb befragt. Die Items waren wie folgt formuliert:

- Die mit den Aufgaben abgefragten Inhalte decken sich in weiten Teilen mit Inhalten, die wir in der Berufsschule vermittelt bekommen.
- Die mit den Aufgaben abgefragten Inhalte sind für die Berufspraxis in meinem Ausbildungsbetrieb von hoher Relevanz.

Die Antworten (vgl. Abbildung 45) zeigen insgesamt, dass für die Probanden augenscheinlich kein Zusammenhang zwischen den Inhalten des Tests und den Inhalten der Berufsschule bzw. der Berufspraxis im Ausbildungsbetrieb besteht. 86% sehen dabei keinen Bezug zur Berufsschule und 87,32% keinen zu den berufspraktischen Inhalten, die sie im Ausbildungsbetrieb erfahren.

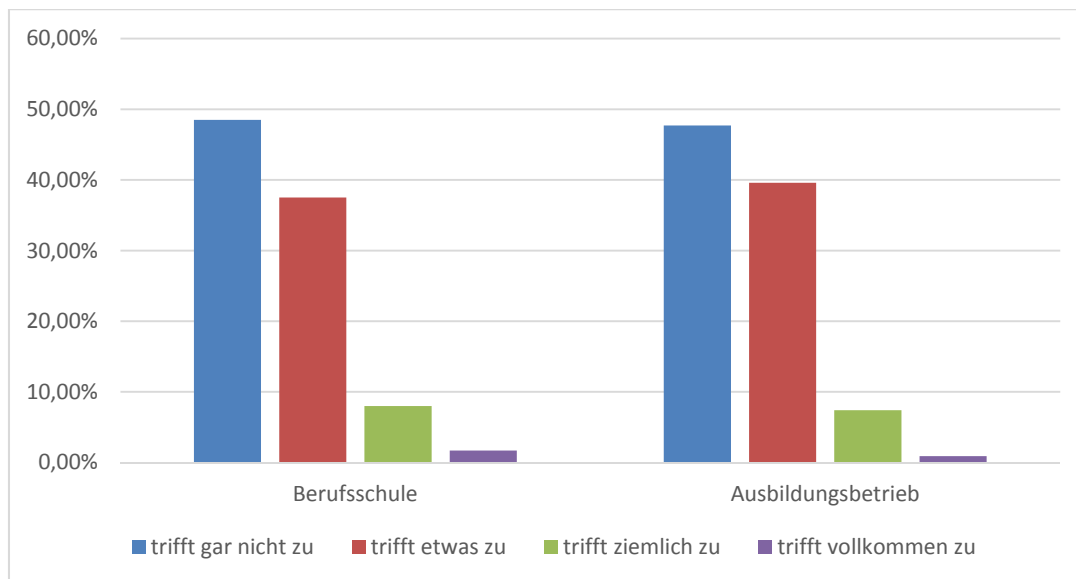


Abbildung 45 – Antwortverhalten Relevanz der Inhalte

Hier wäre es interessant weitere Nachfragen stellen zu können, bspw. welche Inhalte aus dem Bereich Lesen aus Schülersicht denn vermittelt werden bzw. für die Berufspraxis eine hohe Relevanz besitzen. Die Items wurden, wie bereits ausgeführt, nicht speziell für einen Beruf oder eine Berufsgruppe konstruiert, d.h. sie enthalten keine speziellen fachlichen Inhalte eines der getesteten Berufe, stattdessen wurde versucht, die in den verschiedenen Berufen gestellten lesebezogenen Anforderungssituationen abzubilden (vgl. theoretische Herleitung in Kapitel 2.1.2 und Umsetzung in Kapitel 4.1.4). Fraglich ist an dieser Stelle, ob die Schüler von den konkreten Inhalten der Items abstrahieren können und evtl. nach diesbezüglichen Hinweisen die Konstruktähnlichkeit der Items mit arbeitsplatzbezogenen Leseaufgaben erkennen und ggf. bestätigen würden.

5.2. Beantwortung der Forschungsfragen & Reflexion des Untersuchten

Innerhalb dieser Arbeit konnte an Hand vorhergehender theoretischer und empirischer Arbeiten aufgezeigt werden, dass einerseits gute Lesefähigkeit notwendig für beruflichen Erfolg und gesellschaftlichen Einfluss ist, dass aber andererseits sowohl Leseanlässe, Lesematerial und Leseprozesse im Beruf noch wenig erforscht sind. Festgehalten werden kann, dass sich beruflich-betriebliches Lesen durch spezielle Eigenschaften, Anforderungen und Materialien auszeichnet – es könnte also als eine spezifische Facette des Lesens bezeichnet werden. Diese Facette ist vor allem dadurch gekennzeichnet, dass es eine Form des Lesens darstellt, welche zentral durch eine Handlungsintention charakterisiert ist im Gegensatz bspw. zum studierenden Lesen, welches durch eine Lernintention gekennzeichnet ist. Handlungsintention meint, dass die zentrale Funktion des Lesens das Umsetzen des soeben gelesenen ist (bspw. die Bedienung einer Maschine), wozu eine mentale Modellierung der Handlung vorab erforderlich scheint und offenkundig wird, dass es eher unter Handlungsdruck geschieht. Auch scheint in diesem Kontext folgerichtig, dass es sich bei dem Lesematerial zumeist um Gebrauchstexte handelt, statt bspw. um Lerntexte. Es geht sodann beim Lesen im Arbeitskontext nicht um ein umfassendes Erschließen von eher unbekannten Lerninhalten, wie es vielfach im Schul- und Universitätskontext zu finden ist, sondern um eine zielgerichtete Informationssuche zu prinzipiell vertrauten Sachverhalten. Diese Facette des Lesens kann als „funktionales Lesen“ bezeichnet werden (Ziegler et al. 2012).

Lesekompetenz allgemein wird in neueren Definitionen verstanden als die Fähigkeit, multiple Darstellungen in Dokumenten, die Texte, Bilder, Diagramme, Tabellen oder andere Arten externer Repräsentationen enthalten können, zu verstehen und zu nutzen (OECD 1995; Artelt et al. 2001; OECD 2001, 2005, 2012). Hier wird zentral eine Entwicklung in der Forschung der letzten Jahre berücksichtigt: Gelesen werden nicht nur symbolbasierte kontinuierliche Texte (sogenannte deskriptionale Formate) sondern auch vielfach analoge, diskontinuierliche, bildhafte Strukturen wie bspw. Graphen (depiktionale Formate) und Listen (hybride Formate die sowohl analoge wie symbolische Strukturen beinhalten).

Der Prozess des Lesens selbst ist auf den verschiedenen Niveaus (Buchstaben-, Wort-, Satz-, Textebene) weitestgehend gut bekannt, wenngleich es über die letzte Stufe – die mentale Modellierung des Sachverhalts, der somit letzte Schritt des eigentlichen Verstehens – bis heute noch Unstimmigkeiten bezüglich seiner Obligatorik und Fakultativität gibt (Kintsch 1977; Kintsch/van Dijk 1978; Johnson-Laird 1983; Kieras/Bovair 1984; Perrig/Kintsch 1985; Garnham/Oakhill 1996; Friedman/Miyake 2000; Kintsch 2004). Parallel dazu ist die analoge Erforschung des Bildverstehens noch weniger weit fortgeschritten. Obzwar in den letzten Jahren eine starke Entwicklung zu beobachten ist, sind Einzelheiten des Prozesses gerade unter kognitiv informationsverarbeitender Perspektive noch detaillierter auszuarbeiten.

Die gemeinsame Betrachtung von Text und Bild hat vorrangig zur Entstehung zweier theoretischer Schulen geführt: Die „Dual Coding Theory“ (Paivio 1990; Clark/Paivio 1991; Sadoski/Paivio 2001, 2004) und das integrative Text-Bildverstehen (Schnotz/Bannert 1999; Schnotz 2003,

2005, 2006a; Horz/Schnotz 2010). Trotz wesentlicher Unterschiede gehen sie d'accord in der Ansicht, dass der Wechsel von der Text- zur bildlichen Repräsentation oder umgekehrt eine kognitive Leistung ist, die als aufwendig bezeichnet werden kann und die Konstruktion eines Mentalen Modells voraussetzt. Dies ist zentral für die Diagnostik – die Forderung eines Moduswechsels scheint die Abprüfung der Konstruktion eines Mentalen Modells am besten zu ermöglichen.

Verstehendes Lesen ist kein völlig automatisierter Prozess, dessen Erfolg quasi garantiert ist. Es sind seitens des Lesers beträchtliche kognitive Ressourcen aufzubringen, um eine mentale Repräsentation des mitgeteilten Sachverhalts zu etablieren. Das rührt daher, dass einerseits die Kommunikationssituation beim Lesen – Stichwort: zerdehnte Kommunikation – eine völlig andere ist als beim Zuhören, andererseits aber auch daher, dass Texte innerhalb ihrer Form und Struktur auf jeder Ebene verschiedene Parameter aufweisen, die es leichter oder schwerer machen können den Sinn zu entnehmen. Das reicht von Wortbedeutungen über Satzkonstruktionen bis hin zu den inhaltlichen Themen an sich.

Einer der zentralen Aspekte beim Lesen – die Intention des Lesers – und ihr Einfluss ist bisher kaum weiter erforscht. Es gibt einige Hinweise bezüglich ihrer Relevanz sowohl für das Leseergebnis, wie auch für die Ausgestaltung des Leseprozesses (Schmalhofer/Glavanov 1986), jedoch finden diese in der Leseforschung kaum Beachtung. Lesen lässt sich, wie gezeigt, sehr gut und eindeutig über die Intention des Lesers klassifizieren. So gut wie alle anderen am Prozess beteiligten Faktoren (Material, Verarbeitungstiefe, Art der Erfolgskontrolle, Leseziel...) scheinen von der jeweiligen Leserintention abzuhängen.

Um verschiedene Aspekte des beruflichen Lesens in einem diagnostischen Instrument zu vereinen und abbilden zu können, scheint sowohl das Modell der funktionalen Lesekompetenz als auch die Verwendung der Annahme von mentalen Modellen plausibel. Auf dieser Basis wurde im Rahmen eines Forschungsprojekts Items konstruiert, kalibriert und pilotiert. Besondere Berücksichtigung bei der Analyse fand dabei sowohl das Präsentationsformat (depiktional, hybrid, deskriptional) als auch der kognitive Anforderungsbereich (identifizieren, integrieren, generieren) sowie der Präsentationsmodus (Papier vs. PC). Ferner wurden die Items auf Basis von theoretisch hergeleiteten Aspekten aus der Textverständlichkeitsforschung geratet um eine a priori Einschätzung der Schwierigkeit zu gewinnen.

Zentrale Ergebnisse sind:

Die kognitiven Anforderungsbereiche wie identifizieren, integrieren oder generieren weisen bei den vorliegenden Items unterschiedliche Schwierigkeiten auf, wovon sich einige auch als anteilig erklärender Faktor ermitteln ließen (vgl. Kapitel 0). Somit kann die *erste Forschungsfrage* positiv beantwortet werden.

Es zeigte sich, dass die Verarbeitung der verschiedenen Repräsentationsformate keine signifikanten Unterschiede im Anspruch begründen (vgl. Kapitel 0). Damit kann die *zweite Forschungsfrage* wiederum nur negativ beantwortet werden.

Bezüglich der Untersuchung des Unterschieds zwischen papierbasierter und computerbasierter Testung (*dritte Forschungsfrage*) kann festgehalten werden, dass es bedeutsame Unterschiede hinsichtlich der Beantwortung der Items gibt, so dass die beiden Verfahren bezüglich der Schwierigkeitsbestimmung als nicht äquivalent bezeichnet werden müssen und anzunehmen ist, dass auch die kognitiven Prozesse zumindest als andere Facetten des selben Konstrukts betrachtet werden sollten. Dabei konnte auch festgestellt werden, dass der Präsentationsmodus ggf. auch durch die damit verbundenen Rahmenbedingungen (Reihenfolge, notwendige Anzahl der Fragen, motivationale Aspekte...) einen Einfluss auf das Antwortverhalten der Probanden zu zeigen scheinen (vgl. Kapitel 0).

Hinsichtlich der *vierten Forschungsfrage* (vgl. Kapitel 0) kann summiert werden, dass die Interraterreliabilität zwar nach gängigen Konventionen ausreichende Ergebnisse erzielt, diese jedoch, um eine konkrete Vorabeeschätzung von Items zu gewährleisten zu starke Varianzen aufweisen und insgesamt wenig belastbar sind (vgl. Kapitel 0).

Bei einer weiteren Analyse der schwierigkeitsbestimmenden Merkmale (vgl. Kapitel 0) haben sich einige als augenscheinlich eher gute Prädiktoren für eine Steigerung der Anforderungen erwiesen (Inferenzbildung, Konstruktion eines Mentalen Modells, Platzierung der relevanten Informationen). Dabei fällt auf, dass es sich vor allem um solche handelt, die in gewisse Weise als bestimmend für die kognitiven Anforderungen betrachtet werden können. Besonders einflussreich und empirisch aussagekräftig erwies sich der Faktor „Konstruktion eines mentalen Modells“. Wohingegen einige andere Faktoren, die in der Forschung durchaus als schwierigkeitsbestimmende Merkmale identifiziert werden konnten (Antwortformat, Anzahl irrelevanter Informationen), sich für diese Items nicht als aussagekräftig erwiesen haben. So dass die *fünfte Forschungsfrage* tendenziell abschlägig beschieden werden muss.

Der Vergleich der vorgenommen kriteriengeleiteten Einschätzung der Schwierigkeit mit den empirisch ermittelten Lösungswahrscheinlichkeiten zeigt, dass sich das Rating als nicht optimal erweist, da es nur geringe Übereinstimmungen der beiden Verfahren zur Schwierigkeitsbestimmung der Items gibt und diese möglicherweise zufällig zustande gekommen sind (vgl. Kapitel 4.2.5), so dass auch die *sechste Forschungsfrage* negativ beschieden werden muss.

Weitere Forschungsmöglichkeiten – bspw. eine genauere Erforschung der schwierigkeitsbezogenen Parameter – und die Anschlussfähigkeit dieser Befunde sind zahlreich und scheinen notwendig, daher wird auch auf diese kurz im nächsten Kapitel eingegangen.

5.3. Ausblick und Anschlussforschung

Anstoß, Hintergrund und zentraler Anschluss der gesamten Arbeit sind fünf elementare Erkenntnisse:

1. Lesen ist einer der Schlüssel für wirtschaftlichen Erfolg und gesellschaftliche Partizipation in der westlichen Welt.

2. Ein erheblicher Anteil der Jugendlichen erreicht im Bildungssystem nach gängigen Testverfahren nicht die dafür notwendigen Kompetenzen.
3. Elaborierte Lesefördermethoden zeigen kaum oder nicht die gewünschten Effekte, um den unter Punkt 2 angesprochenen Zustand zu beheben und erweisen sich nicht als funktional für berufliche Kontexte.
4. Lesen im Beruf scheint sich augenscheinlich und nach ersten Analysen grundlegend und in wichtigen Punkten (Anforderungskontext) von schulischen und freizeitlichen Leseanlässen zu unterscheiden (Lesen, um zu handeln vs. lesen, um zu lernen vs. lesen, um sich zu unterhalten).
5. Es gibt zurzeit kein elaboriertes Modell von Punkt 3 und dementsprechend kein adäquates Diagnostikinstrument zur Diagnose der konkreten Probleme von Punkt 2 unter Berücksichtigung der Spezifikation von Punkt 4.

Um aber eine erfolgreiche Trainingsmethode anbieten zu können, gilt es den Bogen zum ersten Zitat dieser Arbeit zu schlagen: *Am Anfang steht die Diagnose und jede weitere Handlung ist weitgehend von ebendieser Diagnose bestimmt* (Sydenham). Es ist noch zu wenig über die zentralen Einflussgrößen auf den Leseprozess bekannt, um ein optimales Förderprogramm entwickeln zu können. Diese Arbeit hat einen Beitrag theoretischer Art geleistet, welcher den möglichen Einfluss der Leserintention herausstellt und einen Nachweis empirischer Art, welcher zeigt, dass der Einfluss des Repräsentationsformats schwierig zu untersuchen ist und nach der hier erhobene Datenlage keinen Einfluss auf die Schwierigkeit von Items zu haben scheint. Die belegt den generell geringen Einfluss von textspezifischen Aspekten, gegenüber den situations- und leserspezifischen Voraussetzungen (vgl. Kap. 3.1.5).

Trotz der dargelegten Resultate bleibt auf Grund des Untersuchungsdesigns und der Itemkonstruktion unklar, ob die Dechiffrierungskompetenzen von Texten und Bildern grundsätzlich verschiedene Dimensionen sind oder ob sie alle zufriedenstellend unter Lesefähigkeit subsummiert werden können. Dazu wäre dringend weitere empirische Forschung unter Einbezug von structural equation modelling wünschenswert. Ferner scheinen weiterführende ggf. auch experimentelle Untersuchungen in Bezug auf die Bedeutung der Leseintention notwendig zu sein. Hier ist unter anderem noch unklar, wie groß der Einfluss auf die kognitiven Prozesse ist und in wie weit sich dies auf das Leseergebnis auswirkt. Auch an dieser Stelle scheinen Dimensionsanalysen weitere Erkenntnisse zu bringen. Erste für diese Arbeit durchgeführte Analysen geben zwar Hinweise, scheitern jedoch an fehlender Itemmenge und erfordern eine grundlegende innovative Itemkonstruktion. Insgesamt kann die Itementwicklung noch nicht als zufriedenstellend beurteilt werden. Ihre Deckung mit dem beruflichen Kontext ist noch zu gering und die Induktion einer Handlungsintention behält eher simulativen Charakter, so dass die Abprüfung einer

mental en Modellierung sicherlich optimiert werden könnte. Es zeigt sich aber auch großes Potential bezüglich ihrer Weiterentwicklung – insbesondere beim Abzielen auf auszuführende Handlungen bietet die computerbasierte Testung noch ungenutzte Möglichkeiten.

Bei der Entwicklung von zukünftigen Items scheint es überaus ratsam zu sein, von Single-Choice-Antwortformaten und ein-Wort-Antworten Abstand zu nehmen. Hier bietet die Untersuchung kognitiver Verarbeitungs- und Verstehensprozesse durch Text-Bild-Transformationen noch einige ungenutzte, aber interessante, Möglichkeiten.

Zeitnah geplant ist weiterhin die Untersuchung der vorliegenden ausgefüllten Testhefte auf „Lesespuren“, um die abgelaufenen Prozess rekonstruieren und evtl. aufgetreten Fehler analysieren zu können, um so ein erweitertes Verständnis insbesondere für die Verarbeitung von depiktionalen Strukturen zu gewinnen.

Schlussendlich sollte an die hier erfolgte und in Zukunft weiterzuführende Grundlagenforschung ein Praxistransfer anschließen. Damit ist gemeint, dass alle diese eher auf der theoretisch-erklärenden Seite zu verortenden Erkenntnisse in die Konstruktion und Evaluation eines zuverlässigen und validen beruflichen Lesetests eingehen und im Endeffekt in die Entwicklung eines beruflichen Leseförderprogramms einfließen sollten, um die eingangs angesprochenen Probleme angehen zu können.

6. Literatur

- Adams, Marilyn/Collins, Adam (1977): A schema-theoretic view of reading. Technical Report No. 32.
- Adams, Raymond: Scaling PISA Cognitive Data. In: Adams, Raymond/Wu, Margaret (Hrsg.): PISA 2000. Technical Report. S. 99–108.
- Adams, Raymond/Khoo, Siek-Toon (1996): Quest: the interactive test analysis system. Melbourne: Australian Council for Educational Research.
- Adam-Schwebe, Stefanie/Gold, Andreas/Souvignier, Elmar (2009): Der Frankfurter Leseverständnistest 5-6 (FLVT 5-6). In: Lenhard, Wolfgang/Schneider, Wolfgang (Hrsg.): Diagnostik und Förderung des Leseverständnisses. Göttingen: Hogrefe. S. 113–130.
- Adamzik, Kirsten (2004): Textlinguistik. Eine einführende Darstellung. Tübingen.
- Aebli, Hans (1980): Denken: das Ordnen des Tuns. Band I: Kognitive Aspekte der Handlungstheorie. Stuttgart: Klett Cotta.
- Aebli, Hans (1983): Zwölf Grundformen des Lehrens. Stuttgart.
- Alich, Holger (2011): Die Schweiz kämpft mit der eigenen Währung <http://www.handelsblatt.com/politik/international/euro-schuldenkrise-die-schweiz-kaempft-mit-der-eigenen-waehrung/4508062.html>. Zugriff am 6.1.2014.
- Andersen, Christiane (2009): Instruktion und Multimedialität. Decodierungsprozesse zwischen verbalem und visuellem Teiltex in deutschen und schwedischen Gebrauchsanweisungen. In: Ditlevsen, Marianne/Kastberg, Peter/Andersen, Christiane (Hrsg.): Sind Gebrauchsanleitungen zu gebrauchen? Linguistische und kommunikativ-pragmatische Studien zu skandinavischen und deutschen Instruktionstexten. Tübingen: Gunter Narr Verlag (Europäische Studien zur Textlinguistik).
- Artelt, Cordula/McElvany, Nele/Christmann, Ursula/Richter, Tobias/Groeben, Norbert/Köster, Juliane/Schneider, Wolfgang/Stanat, Petra/Ostermeier, Christian/Schiefele, Ulrich/Valtin, Renate/Ring, Klaus (2007): Förderung von Lesekompetenz. Expertise. Bonn, Berlin.
- Artelt, Cordula/Stanat, Petra (2010): Leistungen von Schülerinnen und Schülern im internationalen Vergleich. Die PISA Studie. In: Spiel, Christiane/Schober, Barbara/Wagner, Petra/Reimann, Ralph (Hrsg.): Bildungspsychologie. Göttingen: Hogrefe. S. 352–356.
- Artelt, Cordula/Stanat, Petra/Schneider, Wolfgang/Schiefele, Ulrich (2001): Lesekompetenz. Testkonzeption und Ergebnisse. In: Baumert, Jürgen/Klieme, Eckhard/Neubrand, Michael/Prenzel, Manfred/Stanat, Petra/Schiefele, Ulrich/Schneider, Wolfgang/Tillmann, Klaus-Jürgen/Weiß, Manfred (Hrsg.): Pisa 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske & Budrych. S. 69–101.
- Artelt, Cordula/Stanat, Petra/Schneider, Wolfgang/Schiefele, Ulrich/Lehmann, Rainer (2004): Die PISA-Studie zu Lesekompetenz. Überblick und weiterführende Analysen. In: Schiefele, Ulrich/Artelt, Cordula/Schneider, Wolfgang/Stanat, Petra (Hrsg.): Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS Verlag für Sozialwissenschaften. S. 139–168.

Asseburg, Regine (2011): Leistungsbereitschaft in Testsituationen. Motivation zur Bearbeitung adaptiver und nicht-adaptiver Leistungstests. Marburg: tectum.

Auer, Michaela/Gruber, Gabriele/Mayringer, Heinz/Wimmer, Heinz (2005a): SLS 5-8 Salzburger Lese-Screening für die Klassenstufen 5-8 - Manual. Bern: Hans Huber Verlag.

Auer, Michaela/Gruber, Hans/Mayringer, Heinz/Wimmer, Heinz (2005b): Salzburger Lese-Screening für die Klassenstufen 5–8. Göttingen: Hogrefe Verlag.

Auslegung durch die Berufsschulen in Hessen (2008): Rahmenlehrplan für den Ausbildungsberuf medizinische Fachangestellte/r. Endergebnis der Beratungen der Arbeitsgruppen Süd-, Mittel- und Nordhessen.

Aust, Hugo (1983): Lesen. Überlegungen zum sprachlichen Verstehen. Tübingen: Max Niemeyer Verlag.

Baddeley, Alan (1986): Working Memory. Oxford: Clarendon Press.

Baddeley, Alan/Logie, Robert (1999): Working Memory. The multiple component model. In: Miyake, Akira/Shah, P. (Hrsg.): Models of working memory. Mechanisms of active maintenance and executive control. New York: Cambridge University Press. S. 28–61.

Bahrack, Harry/Wellman, Cara/Hall, Lynda (1988): The effect of language schemas on learning and retention. In: Gruneberg, Michael/Morris, Peter/Sykes, Robert (Hrsg.): Practical aspects of memory. Current research and issues. Volume 2. Clinical and educational implications. Chichester: John Wiley & Sons. S. 385–389.

Ballstaedt, Steffen-Peter (1995): Bildverstehen und Sprache. In: Spillner, Bernd (Hrsg.): Sprache: Verstehen und Verständlichkeit. Kongressbeiträge zur 25. Jahrestagung der Gesellschaft für angewandte Linguistik GAL e.V. Frankfurt am Main: PETER LANG (Forum Angewandte Linguistik). S. 63–70.

Ballstaedt, Steffen-Peter (1996): Bildverstehen, Bildverständlichkeit. Ein Forschungsüberblick unter Anwendungsperspektive. In: Krings, Hans (Hrsg.): Wissenschaftliche Grundlagen der technischen Kommunikation. Tübingen: Narr. S. 191–233.

Ballstaedt, Steffen-Peter/Mandl, Heinz/Schnotz, Wolfgang/Tergen, Sigmar-Olaf (1981): Texte verstehen, Texte gestalten. München, Wien, Baltimore: Urban & Schwarzenberg.

Bartlett, James/Till, Robert/Levy, Julie (1980): Retrieval characteristics of complex pictures. Effects of verbal encoding. In: Journal of Verbal Learning and Verbal Behavior 4 (19) 1980. S. 430–449.

Baumann, Katharina/Siemon, Jens (2013): Wie viel schriftsprachliche Fähigkeit ist für eine erfolgreiche Berufsausbildung erforderlich? In: Die berufsbildende Schule 10 (65) 2013. S. 285–292.

Baumert, Jürgen/Klieme, Eckhard/Neubrand, Michael/Prenzel, Manfred/Stanat, Petra/Schiefele, Ulrich/Schneider, Wolfgang/Tillmann, Klaus-Jürgen/Weiß, Manfred (Hrsg.) (2001a): Pisa 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske & Budrych.

Baumert, Jürgen/Stanat, Petra/Demmrich, Anke (2001b): PISA 2000. Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Baumert, Jürgen/Klieme, Eckhard/Neubrand, Michael/Prenzel, Manfred/Stanat, Petra/Schiefele, Ulrich/Schneider,

- Wolfgang/Tillmann, Klaus-Jürgen/Weiß, Manfred (Hrsg.): Pisa 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske & Budrych. S. 15–68.
- Beaton, Albert/Allen, Nancy (1992): Interpreting Scales Through Scale Anchoring. In: Journal of Educational Statistics 2 (17) 1992. S. 191–204.
- Becker-Mrotzek, Michael/Drommler, Rebecca (2006): Texte lesen. In: Becker-Mrotzek, Michael/Kusch, Erhard/Wehnert, Bernd (Hrsg.): Leseförderung in der Berufsbildung. Kölner Beiträge zur Sprachdidaktik. Duisburg: Gilles & Francke. S. 17–44.
- Becker-Mrotzek, Michael/Kusch, Erhard/Wehnert, Bernd (Hrsg.) (2006): Leseförderung in der Berufsbildung. Kölner Beiträge zur Sprachdidaktik. Duisburg: Gilles & Francke.
- Bernhardt, Raphael/Frey, Andreas (2013): Konstruktion einer computerbasierten adaptiven Testbatterie zur Messung von Kompetenzen in Mathematik, Lesen und Naturwissenschaften von Schülerinnen und Schülern beruflicher Schulen. Vortrag auf der 1. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF). Kiel.
- Bertin, Jacques (1974): Graphische Semiologie. Diagramme, Netze, Karten. Berlin: Gruyter Verlag.
- Beschluss der Kultusministerkonferenz (2001): Rahmenlehrplan für den Ausbildungsberuf Industriekaufmann/Industriekauffrau.
- Beschluss der Kultusministerkonferenz (2003a): Rahmenlehrplan für den Ausbildungsberuf Elektroniker für Automatisierungstechnik.
- Beschluss der Kultusministerkonferenz (2003b): Rahmenlehrplan für den Ausbildungsberuf Elektroniker für Betriebstechnik.
- Beschluss der Kultusministerkonferenz (2003c): Rahmenlehrplan für den Ausbildungsberuf Kraftfahrzeugmechatroniker/Kraftfahrzeugmechatronikerin.
- Beschluss der Kultusministerkonferenz (2004): Rahmenlehrplan für den Ausbildungsberuf Kaufmann für Spedition und Logistikdienstleistung/ Kauffrau für Spedition und Logistikdienstleistung.
- Beschluss der Kultusministerkonferenz (2005): Rahmenlehrplan für den Ausbildungsberuf Medizinischer Fachangestellter/Medizinische Fachangestellte.
- Bhola, Harban (1995): Functional literacy, workplace literacy and technical and vocational education. Interfaces and policy Perspectives. Paris: UNESCO.
- Biedenbach, Wyrola (2006): Der Modellversuch "Vocational Literacy (VoLi) - Methodische und sprachliche Kompetenzen in der beruflichen Bildung. Konzeption - Erfahrung - bisherige Ergebnisse. In: Efinger, Christian/Janich, Nina (Hrsg.): Förderung der Berufsbezogenen Sprachkompetenz. Befunde und Perspektiven. Paderborn: Eusl. S. 15–31.
- Birkerts, Sven (1994): The Gutenberg Elegies. The Fate of Reading in an Electronic Age. Boston, London: Faber and Faber.
- Bobrow, Daniel (1975): Some principles of memory schemata. In: Bobrow, Daniel/Collins, Allan (Hrsg.): Representation and Understanding. Studies in Cognitive Science. London: Academic Press. S. 131–149.

- Bock, Michael (1978): Wort-, Satz-, Textverarbeitung. Stuttgart, Berlin, Köln, Mainz: Kohlhammer.
- Bohlinger, Sandra (2007/2008): Kompetenzen als Kernelement des Europäischen Qualifikationsrahmens. In: Europäische Zeitschrift für Berufsbildung 1/3 (42/43) 2007/2008. S. 112–130.
- Bond, Trevor/Fox, Christine (2001): Applying the Rasch Model. Fundamental Measurement in the Human Sciences. Mahwah: Lawrence Erlbaum.
- Bonfadelli, Heinz (1980): Neue Fragestellungen in der Wirkungsforschung. Zur Hypothese der wachsenden Wissensklüft. In: Rundfunk und Fernsehen 2 (28) 1980. S. 173–193.
- Borneto, Carlo/Cortelazzo, Michelle (1988): Textstruktur und Verständlichkeit der Gebrauchsanweisung für elektrische Haushaltsgeräte. In: Bungarten, Theo (Hrsg.): Sprache und Information in Wirtschaft und Gesellschaft. Todstedt: Attikon Verlag. S. 206–218.
- Bortz, Jürgen/Lienert, Gustav/Boehnke, Klaus (2000): Verteilungsfreie Methoden in der Biostatistik. 2. Aufl. Berlin: Springer Verlag.
- Bos, Wilfried/Lankes, Eva-Maria/Schwippert, Knut/Valtin, Renate/Voss, Andreas/Badel, Isolde/Platzmeier, Nike (2003): Lesekompetenzen deutscher Grundschülerinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In: Bos, Wilfried/Lankes, Eva-Maria/Prenzel, Manfred/Schwippert, Knut/Walther, Gerd/Valtin, Renate (Hrsg.): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann.
- Brinker, Klaus (2005): Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden. Berlin: Erich Schmidt Verlag.
- Britton, Bruce/Graesser, Arthur (Hrsg.) (1996): Models of understanding text. Mahwah, NJ: Erlbaum.
- Brosius, Felix (1998): SPSS 8. Professionelle Statistik unter Windows. Bonn: International Thomson Publishing.
- Brünner, Gisela (1998): Fachkommunikation im Betrieb. am Beispiel der Stadtwerke einer Großstadt. In: Hoffmann, Lothar/Kalverkämper, Hartwig/Wiegand, Herbert (Hrsg.): Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft. Berlin, New York: Walter de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft). S. 634–648.
- Buhlmann, Rosemarie/Fearns, Anneliese (2000): Handbuch des Fachsprachenunterrichts. Unter besonderer Berücksichtigung naturwissenschaftlich-technischer Fachsprachen. Tübingen.
- Bühner, Markus (2006): Einführung in die Test- und Fragebogenkonstruktion. 2. Aufl. München: Pearson Studium.
- Bundesagentur für Arbeit (Hrsg.) (2009): Kriterienkatalog zur Ausbildungsreife. Nationaler Pakt für Ausbildung und Fachkräftenachwuchs in Deutschland.
- Bungarten, Theo (1988): Sprache und Information in Wirtschaft und Gesellschaft. In: Bungarten, Theo (Hrsg.): Sprache und Information in Wirtschaft und Gesellschaft. Todstedt: Attikon Verlag. S. 13–22.

- Busse, Dietrich (2000): Textsorten des Bereichs Rechtswesen und Justiz. In: Antos, Gerd/Brinker, Klaus/Heinemann, Wolfgang/Sager, Sven (Hrsg.): Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung. Berlin, New York: De Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft). S. 658–675.
- Cattell, James (1986): The time it takes to see and name objects. In: *Mind* 11 1986. S. 63–65.
- Centre for Educational Research and Innovation (CERI) (1992): Adult Illiteracy and Economic Performance. Paris: Organisation for Economic Co-operation and Development (OECD).
- Christmann, Ursula (1989): Modelle der Textverarbeitung. Textbeschreibung als Textverstehen. Münster: Aschendorff.
- Christmann, Ursula (2000): Aspekte der Textverarbeitungsforschung. In: Brinker, Klaus/Antos, Gerd/Heinemann, Wolfgang/Sager, Sven (Hrsg.): Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung. Berlin, New York: De Gruyter. S. 113–122.
- Christmann, Ursula/Groeben, Norbert (1996): Textverstehen, Textverständlichkeit. Ein Forschungsüberblick unter Anwendungsperspektive. In: Krings, Hans (Hrsg.): Wissenschaftliche Grundlagen der technischen Kommunikation. Tübingen: Narr. S. 129–189.
- Christmann, Ursula/Groeben, Norbert (2006): Psychologie des Lesens. In: Franzmann, Bodo/Hasemann, Klaus/Löffler, Dietrich/Schön, Erich (Hrsg.): Handbuch Lesen 2. Aufl. Baltmannsweiler: Schneider Verlag Hohengehren. S. 145–223.
- Clark, James/Paivio, Allan (1991): Dual Coding Theory and Education. In: *Educational Psychology Review* 3 (3) 1991. S. 149–210.
- Cohen, Jacob (1992): A Power Primer. In: *Psychological Bulletin* 1 (112) 1992. S. 155–159.
- Coltheart, Max (1981): Disorders of Reading and Their Implications for Models of Normal Reading. In: *Visible Language* 15 1981. S. 245–286.
- Dallmann, Gerhard/Preibusch, Wolfgang (1970): Unterrichtsmedien. In: Ingenkamp, K./Parey, E. (Hrsg.): Handbuch der Unterrichtsforschung. Band 2. Weinheim: Beltz. S. 1530–1799.
- Deutscher Bildungsrat (1974): Empfehlungen der Bildungskommission. Zur Neuordnung der Sekundarstufe II. Konzept für eine Verbindung von allgemeinem und beruflichem Lernen. Stuttgart.
- Diehl, William/Mikulecky, Larry (1980): The Nature of Reading at Work. In: *Journal of Reading* 3 1980. S. 221–227.
- Dilger, Bernadette (2004): Kompetenz als Standard der Bildung (von Standards). In: *Wirtschaft und Pädagogik* 36 2004. S. 11–36.
- Drommler, Rebecca/Linnemann, Markus/Becker-Mrotzek, Michael/Haider, Hilde/Stevens, Tobias/Wahlers, Judith (2006): Lesetest für Berufsschüler/innen. LTB-Handbuch. Duisburg: Gilles & Francke.
- Drösser, Christoph (12.06.2003): "Stimmts?". Blitztod im Teich. In: *Die Zeit* 25 12.06.2003.
- Efing, Christian (2006): "Viele sind nicht in der Lage, diese schwarzen Symbole da lebendig zu machen" - Befunde empirischer Erhebungen zur Sprachkompetenz hessischer Berufsschullehrer. In: Efing, Christian/Janich, Nina (Hrsg.): Förderung der Berufsbezogenen Sprachkompetenz. Befunde und Perspektiven. Paderborn: Eusl. S. 33–68.

Efing, Christian (2012): Sprachliche oder kommunikative Fähigkeiten - Was ist der Unterschied und was wird in der Ausbildung verlangt? In: Berufsbildung in Wissenschaft und Praxis 2 2012. S. 6–9.

Efing, Christian (2013): Sprachlich-kommunikative Anforderungen in der betrieblichen Ausbildung. In: Efing, Christian (Hrsg.): Ausbildungsvorbereitung im Deutschunterricht der Sekundarstufe I. Die sprachlich-kommunikativen Facetten von "Ausbildungsfähigkeit". Frankfurt am Main: PETER LANG (Wissen - Kompetenz - Text). S. 123–145.

Ehlich, Konrad (1994): Verweisung und Kohärenz in Bedienungsanleitungen. Einige Aspekte der Verständlichkeit von Texten. In: Ehlich, Konrad/Noack, Claus/Scheiter, Susanne (Hrsg.): Instruktion von Text und Diskurs. Zur Linguistik "Technischer Texte". Opladen: Westdeutscher Verlag. S. 116–149.

Elwert, Georg (2001): Societal Literacy. Writing Culture and Development. In: Olson, David/Torrance, Nancy (Hrsg.): The making of Literate Societies. Malden, Oxford: Blackwell Publishers Inc.. S. 54–67.

Engelkamp, Johannes (1990): Das menschliche Gedächtnis. Das Erinnern von Sprache, Bildern und Handlungen. Göttingen: Hogrefe.

Europäische Kommission (2000): Bericht über die Qualität der schulischen Bildung in Europa. Sechzehn Qualitätsindikatoren. Bericht auf der Grundlage der Ergebnisse des Arbeitsausschusses „Qualitätsindikatoren“. Luxemburg: Amt für amtliche Veröffentlichungen der europäischen Gemeinschaften.

Feldmann, Laurie (1991): The Contribution of Morphology to Word Recognition. In: Psychological Research 53 1991. S. 33–41.

Fickermann, Ingeborg (1994): Mündliche und schriftliche Instruktion. In: Ehlich, Konrad/Noack, Claus/Scheiter, Susanne (Hrsg.): Instruktion von Text und Diskurs. Zur Linguistik "Technischer Texte". Opladen: Westdeutscher Verlag. S. 79–99.

Fluck, Hans-Rüdiger (1992): Didaktik der Fachsprachen. Aufgaben und Arbeitsfelder, Konzepte und Perspektiven im Sprachbereich Deutsch. Tübingen.

Fodor, Jerry (1983): The modularity of mind. Cambridge.

Fotheringham, Heinz (1982): Gebrauchsinformationen und -anweisungen aus juristischer Sicht. In: Grosse, Siegfried/Mentrup, Wolfgang (Hrsg.): Anweisungstexte. Tübingen: Gunter Narr Verlag (Forschungsbericht des Instituts für deutsche Sprache). S. 69–73.

Fraas, Claudia (1998): Lexikalisch-semantische Eigenschaften von Fachsprachen. In: Hoffmann, Lothar/Kalverkämper, Hartwig/Wiegand, Herbert (Hrsg.): Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft. Berlin, New York: Walter de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft). S. 428–438.

Frey, Andreas (2004): Die Kompetenzstruktur von Studierenden des Lehrerberufs. Eine internationale Studie. In: Zeitschrift für Pädagogik 6 (50) 2004. S. 903–925.

Frey, Andreas (2012): Adaptives Testen. In: Moosbrugger, Helfried/Kelava, Augustin (Hrsg.): Testtheorie und Fragebogenkonstruktion 2. Aufl. Berlin Heidelberg: Springer Verlag. S. 261–278.

- Frey, Andreas/Ehmke, Timo (2007): Hypothetischer Einsatz adaptiven Testens bei der Überprüfung von Bildungsstandards. In: Prenzel, Manfred/Gogolin, Ingrid/Krüger, Heinz-Hermann (Hrsg.): Kompetenzdiagnostik. Wiesbaden: VS Verlag (Zeitschrift für Erziehungswissenschaft). S. 169–184.
- Frey, Andreas/Seitz, Nicki-Nils (2010): Multidimensionale adaptive Kompetenzdiagnostik. Ergebnisse zur Messeffizienz. In: Zeitschrift für Pädagogik 56 2010. S. 40–51.
- Friedman, Naomi/Miyake, Akira (2000): Differential Roles for Visuospatial and Verbal Working Memory in Situation Model Construction. In: Journal of Experimental Psychology 1 2000. S. 61–83.
- Fritz, Angela (1991): Lesen im Medienumfeld. Eine Studie zur Entwicklung und zum Verhalten von Lesern in der Mediengesellschaft auf der Basis von Sekundäranalysen zur Studie "Kommunikationsverhalten und Medien" im Auftrag der Bertelsmann Stiftung. Gütersloh: Verlag Bertelsmann Stiftung.
- Fritz, Angela/Suess, Alexandra (1986): Lesen. Die Bedeutung der Kulturtechnik Lesen für den gesellschaftlichen Kommunikationsprozeß. Konstanz: Universitätsverlag Konstanz.
- Fuchshuber-Weiß, Elisabeth (1985): Gebrauchsanweisungen für Phonogeräte. Routine, Irritation und Veränderung. Gewohnheitswissen und Textverständnis bei Dreizehnjährigen. In: Der Deutschunterricht 2 (37) 1985. S. 6–17.
- Garbe, Christine (2009): Lesekompetenz. In: Garbe, Christine/Holle, Karl/Jesch, Tatjana (Hrsg.): Texte lesen. Textverstehen Lesedidaktik Lesesozialisation. Paderborn: Ferdinand Schöningh Verlag.
- Garfield, Jay (Hrsg.) (1989): Modularity in knowledge representation and natural-language understanding. Cambridge.
- Garnham, Alan/Oakhill, Jane (1996): The mental models theory of language comprehension. In: Britton, Bruce/Graesser, Arthur (Hrsg.): Models of understanding text. Mahwah, NJ: Erlbaum. S. 313–339.
- Gartz, Michael/Hüchtermann, Marion/Mrytz, Barbara (1999): Schulabgänger. Was sie können und was sie können müßten. Köln: Deutscher Insitutsverlag.
- Gilomen, Heinz (2003): Desired outcomes. A successful life and a well-functioning society. In: Rychen, Dominique/Salganik, Laura (Hrsg.): Key-Competencies for a successful life and a well-functioning society. Cambridge: Hogrefe + Huber. S. 109–135.
- Gläser, Rosemarie (1990): Fachtextsorten im Englischen. Tübingen.
- Glenberg, Arthur/Kruley, Peter/Langston, William (1994): Analogical processes in comprehension. Simulation of a mental model. In: Gernsbacher, Morton (Hrsg.): Handbook of Psycholinguistics. San Diego: Academic Press. S. 609–640.
- Glenberg, Arthur/Langston, William (1992): Comprehension of illustrated text. Pictures help to build mental models. In: Journal of Memory and Language (31) 1992. S. 123–151.
- Glenberg, Arthur/Meyer, Marion/Lindem, Karen (1987): Mental Models contribute to foregrounding during text comprehension. In: Journal of Memory and Language 26 1987. S. 69–83.

- Gloning, Thomas (1995): Zur Verständlichkeit von Packungsbeilagen. In: Spillner, Bernd (Hrsg.): Sprache: Verstehen und Verständlichkeit. Kongressbeiträge zur 25. Jahrestagung der Gesellschaft für angewandte Linguistik GAL e.V. Frankfurt am Main: PETER LANG (Forum Angewandte Linguistik). S. 44–55.
- Glück, Helmut (2000): Metzler-Lexikon Sprache. 2. Aufl. Stuttgart, Weimar: Metzler Verlag.
- Gnahs, Dieter (2007): Kompetenzen – Erwerb, Erfassung, Instrumente. Bielefeld.
- Gough, Philip (1972): One Second of Reading. In: Kavanagh, James/Mattingly, Ignatius (Hrsg.): Language By Ear And By Eye. The relationships between speech and reading. Cambridge: MIT Press. S. 331–358.
- Gould, John/Grischkowsky, Nancy (1984): Doing the Same Work with Hard Copy and with Cathode-Ray Tube (CRT) Computer Terminals. In: Human Factors: The Journal of the Human Factors and Ergonomic Society 3 (26) 1984. S. 323–337.
- Grabowski, Joachim (Hrsg.) (2014): Sinn und Unsinn von Kompetenzen. Fähigkeitskonzepte im Bereich von Sprache, Medien und Kultur.
- Grimes, Joseph (1984): The thread of discourse. 3. Aufl. Berlin: De Gruyter Mouton.
- Groeben, Norbert (1978): Die Verständlichkeit von Unterrichtstexten. Dimensionen und Kriterien rezeptiver Lernstadien. 2. Aufl. Münster: Aschendorff.
- Groeben, Norbert (1982): Leserpsychologie. Textverständnis - Textverständlichkeit. Münster: Aschendorff.
- Groeben, Norbert (2002): Zur konzeptuellen Struktur des Konstrukts "Lesekompetenz". In: Groeben, Norbert/Hurrelmann, Bettina (Hrsg.): Lesekompetenz. Bedingungen, Dimensionen, Funktionen. Weinheim: Juventa.
- Groeben, Norbert/Hurrelmann, Bettina (Hrsg.) (2002): Lesekompetenz. Bedingungen, Dimensionen, Funktionen. Weinheim: Juventa.
- Groeben, Norbert/Hurrelmann, Bettina (2004): Fazit: Lesen als Schlüsselqualifikation? In: Groeben, Norbert/Hurrelmann, Bettina (Hrsg.): Lesesozialisation in der Mediengesellschaft. Ein Forschungsüberblick. Weinheim, München: Juventa. S. 440–465.
- Gschwendtner, Tobias (2012): Förderung des Leseverständnisses in Benachteiligtenklassen der beruflichen Bildung: Studien zur Implementation und Wirksamkeit von Reciprocal Teaching. Aachen: Shaker Verlag.
- Günther, Hartmut (1988): Schriftliche Sprache. Strukturen geschriebener Wörter und ihre Verarbeitung beim Lesen. Tübingen: M. Niemeyer.
- Hansen, Wilfred/Doring, Richard/Whitlock, Lawrence (1978): Why an examination was slower on-line than on paper. In: International Journal of Man-Machine Studies 10 1978. S. 507–519.
- Hansen-Schirra, Silvia/Neumann, Stella (2004): Linguistische Verständlichmachung in der juristischen Realität. In: Lerch, Kent (Hrsg.): Die Sprache des Rechts. Recht verstehen. Band 1. Berlin: De Gruyter. S. 167–184.
- Hartig, Johannes (2007): Skalierung und Definition von Kompetenzniveaus. In: Beck, Bärbel (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch-Englisch-Schülerleistungen-International). Weinheim, Basel: Beltz. S. 83–99.

- Hartig, Johannes (2008): Kompetenzen als Ergebnisse von Bildungsprozessen. In: Jude, Nina/Hartig, Johannes/Klieme, Eckhard (Hrsg.): Kompetenzfassung in pädagogischen Handlungsfeldern. Theorien, Konzepte und Methoden. Berlin, Bonn. S. 13–24.
- Hartig, Johannes/Klieme, Eckhard (2006): Kompetenz und Kompetenzdiagnostik. In: Schweizer, Karl (Hrsg.): Leistung und Leistungsdiagnostik. Heidelberg: Springer Medizin Verlag. S. 127–141.
- Hartung, Silvia (1992): Die Textsortenproblematik in der deutschen Gegenwartssprache. Beschreibung der Textsuche "Bedienungsanleitung". In: Sommerfeldt, Karl-Ernst (Hrsg.): Vom Satz zum Text. Frankfurt: PETER LANG (Sprache System und Tätigkeit). S. 121–133.
- Hensel, Cornelia (1989): Produktbegleitende Texte. Der Versuch einer Analyse unter illokutionären Aspekt. In: Fleischer, Wolfgang/Große, Rudolf/Lerchner, Gotthard (Hrsg.): Beiträge zur Erforschung der deutschen Sprache. Leipzig. S. 138–157.
- Herrmann, Theo (1990): Sprechen und Sprachverstehen. In: Spada, Hans (Hrsg.): Lehrbuch Allgemeine Psychologie. Bern: Hans Huber Verlag. S. 281–322.
- Hessischen Sozialministerium (2011): Rahmenlehrplan für die schulische und betriebliche Ausbildung. Fachkraft Altenpflege. 2. Aufl.
- Hirst, Russel (1990): Using visual mnemonics to make instructions easier to remember. In: Journal of Technical Writing 20 1990. S. 411–423.
- Hockett, Charles (1958): A course in modern linguistics. New York.
- Hoffmann, Lothar (1985): Kommunikationsmittel Fachsprache. Eine Einführung. 2. Aufl. Tübingen.
- Hoffmann, Lothar (1987): Fachsprachen. Instrument und Objekt. Leipzig.
- Hoffmann, Ludger (1983): Arzneimittel-Gebrauchsinformationen. Struktur, kommunikative Funktionen und Verständlichkeit. In: Deutsche Sprache 11 1983. S. 138–159.
- Holloway, John (1999): Improving the Reading Skills of Adolescents. In: Educational Leadership 2 (57) 1999. S. 80–81.
- Horz, Holger/Schnotz, Wolfgang (2010): Multimedialer Wissenserwerb. In: Spiel, Christiane/Schober, Barbara/Wagner, Petra/Reimann, Ralph (Hrsg.): Bildungspsychologie. Göttingen: Hogrefe. S. 188–192.
- Hurrelmann, Bettina (2008): Modelle und Merkmale der Lesekompetenz. In: Bertschi-Kaufmann, Andrea (Hrsg.): Lesekompetenz, Leseleistung, Leseförderung. Grundlagen, Modelle, Materialien 2. Aufl. Zug: Klett und Balmer/ Klett Kallmeyer. S. 18–28.
- Husfeldt, Vera (2001): Literalität, Bildung und Beschäftigung. Münster: Waxmann.
- Hymowitz, L. (22.01.1981): Employers Take over Where School Failed to Teach the Basics. In: Wall Street Journal 22.01.1981. S. 1.
- Industrie- und Handelskammer Frankfurt am Main (2004): Ausbildungsrahmenplan Kraftfahrzeugmechatroniker.
- Jakobs, Eva-Maria (2006): Texte im Berufsalltag. Schreiben, um verstanden zu werden. In: Blühdorn, Hardarik/Breindl, Eva/Waßer, Ulrich (Hrsg.): Text - Verstehen. Grammatik und darüber hinaus. Berlin, New York: Walter de Gruyter. S. 315–350.

- Jakobs, Eva-Maria (2007): „Das lernt man im Beruf...“. Schreibkompetenz für den Arbeitsplatz. In: Werlen, Erika/Tissot, Fabienne (Hrsg.): Sprachvermittlung in einem mehrsprachigen kommunikationsorientierten Umfeld. Hohengehren: Schneider Verlag (Sprachenlernen konkret). S. 27–42.
- Jakobs, Eva-Maria (2008a): Berufliches Schreiben: Ausbildung, Training, Coaching. Überblick zum Gegenstand. In: Jakobs, Eva-Maria/Lehnen, Katrin (Hrsg.): Berufliches Schreiben. Ausbildung, Training, Coaching. Frankfurt am Main, Berlin, Bern, Brüssel, New York, Oxford, Wien: PETER LANG (Textproduktion und Medium). S. 1–14.
- Jakobs, Eva-Maria (2008b): Textproduktion und Kontext. Domänenspezifisches Schreiben. In: Janich, Nina (Hrsg.): Textlinguistik. 15 Einführungen. Tübingen: Narr. S. 255–379.
- Jakobs, Eva-Maria/Schindler, Kirsten (2006): Wie viel Kommunikation braucht der Ingenieur? Ausbildungsbedarf in technischen Berufen. In: Efing, Christian/Janich, Nina (Hrsg.): Förderung der Berufsbezogenen Sprachkompetenz. Befunde und Perspektiven. Paderborn: Eusl. S. 133–153.
- Jenkins, James (1979): Four points to remember. A tetrahedral model and memory experiments. In: Cermak, L./Craik, F. (Hrsg.): Levels of processing in human memory. Hillsdale, NJ: Lawrence Erlbaum. S. 429–446.
- Johnson-Laird, Philip (1983): Mental models. Towards a cognitive science of language, inference, and consciousness. Cambridge, Mass: Harvard University Press.
- Jones, Stan (1996): Economic dimensions of literacy in Canada. In: Statistics Canada (Hrsg.): Reading the Future. A Portrait of Literacy in Canada. Ottawa. S. 39–55.
- Jordan, Roland (2011): Entwicklung und Validierung eines Testverfahrens zur Ermittlung der Lesekompetenz und des mathematischen Textverständnisses mit empirischer Untersuchung an allgemeinbildenden und berufsbildenden Schulen. Münster: Verlag für Wissenschaftliche Texte und Medien.
- Jude, Nina (23.02.2006): IRT-Skalierung mit ConQuest. Workshop für das Nachwuchsnetzwerk Deutschdidaktik. Hamburg.
- Juhl, Dietrich (1994): Vermittlungsstrategien in Bedienungsanleitungen. In: Ehlich, Konrad/Noack, Claus/Scheiter, Susanne (Hrsg.): Instruktion von Text und Diskurs. Zur Linguistik "Technischer Texte". Opladen: Westdeutscher Verlag. S. 198–204.
- Kaiser, Franz (2012): Sprache - Handwerkszeug kaufmännischer Berufe. In: Berufsbildung in Wissenschaft und Praxis 2 2012. S. 14–17.
- Keimes, Christina (2014): Lesen. Eine empirische Studie zur berufsspezifischen Relevanz des Lesens in gewerblich-technischen Bildungsgängen. Im Druck.
- Keimes, Christina/Rexing, Volker (2011a): Förderung der Lesekompetenz von Berufsschülerinnen und Berufsschülern. Bilanz von Fördermaßnahmen. In: Zeitschrift für Berufs- und Wirtschaftspädagogik 1 2011. S. 77–91.
- Keimes, Christina/Rexing, Volker (2011b): Leseanforderungen im Kontext beruflicher Arbeit im Berufsfeld Bautechnik - empirische Befunde und Konsequenzen für die Lesekompetenzförderung. In: Berufs- und Wirtschaftspädagogik online spezial (5) 2011. S. 1–15.

- Keimes, Christina/Rexing, Volker/Ziegler, Birgit (2011): Leseanforderungen im Kontext beruflicher Arbeit. Empirische Befunde und Konsequenzen für die Entwicklung adressatenspezifischer integrierter Konzepte zur Förderung von Lesestrategien. In: Die berufsbildende Schule 7/8 2011. S. 227–232.
- Kelter, Stephanie (1995): Räumliche Vorstellung und Textverstehen. Neuere Entwicklungen der Theorie mentaler Modelle. In: Spillner, Bernd (Hrsg.): Sprache: Verstehen und Verständlichkeit. Kongressbeiträge zur 25. Jahrestagung der Gesellschaft für angewandte Linguistik GAL e.V. Frankfurt am Main: PETER LANG (Forum Angewandte Linguistik). S. 70–82.
- Kieras, David/Bovair, Susan (1984): The role of a mental model in learning to operate a device. In: Cognitive Science 8 1984. S. 255–273.
- Kintsch, Walter (1974): The representation of meaning in memory. Hillsdale, NJ: Lawrence Erlbaum.
- Kintsch, Walter (1977): Memory and Cognition. New York: John Wiley & Sons.
- Kintsch, Walter (1994): Text comprehension, memory, and learning. In: American Psychologist (49) 1994. S. 294–303.
- Kintsch, Walter (1998): Comprehension. A paradigm for cognition. Cambridge: Cambridge University Press.
- Kintsch, Walter (2004): The Construction-Integration model of text comprehension and its implications for instruction. In: Ruddel/Unrau (Hrsg.): Theoretical models and processes of reading 5. Aufl. Newark. S. 1270–1328.
- Kintsch, Walter/van Dijk, Teun (1978): Toward a model of Text comprehension and Production. In: Psychological Review 5 (85) 1978. S. 363–394.
- Kintsch, Walter/van Dijk, Teun (1983): Strategies of Discourse Comprehension. New York: Academic Press.
- Kirsch, Irwin (1999): Lesekompetenz auf drei Skalen. Definitionen und Ergebnisse. In: Notter, Phillip/Bonerad, Eva-Marie/Stoll, Francois (Hrsg.): Lesen - eine Selbstverständlichkeit? Schweizer Bericht zum "International Adult Literacy Survey". Zürich: Verlag Ruegger. S. 191–234.
- Kirsch, Irwin/de Jong, John/Lafontaine, Dominique/McQueen, Joy/Mendelovits, Juliette/Monseur, Christian (2002a): Lesen kann die Welt verändern. Leistung und Engagement im Ländervergleich. Ergebnisse von Pisa 2000.
- Kirsch, Irwin/Guthrie, John (1977): The concept and measurement of functional literacy. In: Reading Research Quarterly 4 (78) 1977. S. 485–507.
- Kirsch, Irwin/Jungblut, Ann/Jenkins, Lynn/Kolstad, Andrew (2002b): Adult Literacy in America. A First Look at the Findings of the National Adult Literacy Survey. 3. Aufl.
- Kirscher, Barbara (1998): Hilfe - der Text ist zu schwer! Typische sprachliche Stolpersteine in Fachtexten - und wie man sie umgeht. In: Lernchancen 59 1998.
- Klicpera, Christian/Gasteiger-Klicpera, Barbara (1993): Lesen und Schreiben. Entwicklung und Schwierigkeiten. Bern: Hans Huber Verlag.

Klicpera, Christian/Gasteiger-Klicpera, Barbara (1995): Psychologie der Lese- und Schreibschwierigkeiten. Entwicklung, Ursachen, Förderung. 2. Aufl. Weinheim: Beitz Psychologie Verlagsunion.

Klieme, Eckhard (2004): Was sind Kompetenzen und wie lassen sie sich messen? In: Pädagogik 6 2004. S. 10–13.

Klieme, Eckhard/Hartig, Johannes (2007): Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In: Prenzel, Manfred/Gogolin, Ingrid/Krüger, Heinz-Hermann (Hrsg.): Kompetenzdiagnostik. Wiesbaden: VS Verlag (Zeitschrift für Erziehungswissenschaft). S. 11–29.

Klieme, Eckhard/Leutner, Detlev (2005): Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Antrag an die Deutsche Forschungsgemeinschaft auf Einrichtung eines Schwerpunktprogramms.

Körber, Andreas (2007): Grundbegriffe und Konzepte. Bildungsstandards, Kompetenzen und Kompetenzmodelle. In: Körber, Andreas/Schreiber, Waltraud/Schöner, Alexander (Hrsg.): Kompetenzen Historischen Denkens. Ein Strukturmodell als Beitrag zur Kompetenzorientierung in der Geschichtsdidaktik. Neuried: ars una. S. 54–86.

Krahn, Harvey/Lowe, Graham (1999): Literacy in the workplace. In: PERSPECTIVES 1999. S. 28–44.

Kretzschmar, Franziska/Pleimling, Dominique (2011): Nähere Informationen zur Lesestudie: Unterschiedliche Lesegeräte, unterschiedliches Lesen? Studienpapier zum Projekt.

Krings, Hans (1996): Wie viel Wissenschaft brauchen technische Redakteure? Zum Verhältnis von Wissenschaft und Praxis in der technischen Dokumentation. In: Krings, Hans (Hrsg.): Wissenschaftliche Grundlagen der technischen Kommunikation. Tübingen: Narr. S. 5–128.

Küster, Rainer (1982): Pragmalinguistische Aspekte von Anweisungstexten. In: Grosse, Siegfried/Mentrup, Wolfgang (Hrsg.): Anweisungstexte. Tübingen: Gunter Narr Verlag (Forschungsbericht des Instituts für deutsche Sprache). S. 104–133.

Lachmeyer, Simone (2008): Entwicklung und Überprüfung eines Strukturmodells der Diagrammkompetenz für den Biologieunterricht. Disseration: Universität Kiel.

Ladenthin, Volker (2010): Kompetenzorientierung als Indiz pädagogischer Orientierungslosigkeit. In: Vierteljahrsschrift für wissenschaftliche Pädagogik 3 (86) 2010. S. 346–358.

Landis, Richard/Koch, Gary (1977): The Measurement of Observer Agreement for Categorical Data. In: Biometrics 1 (33) 1977. S. 159–174.

Lauf, Edmund (2001): ".96 nach Holsti". Zur Reliabilität von Inhaltsanalysen und deren Darstellung in kommunikationswissenschaftlichen Fachzeitschriften. In: Publizistik 1 (46) 2001. S. 57–68.

Lehmann, Rainer/Hunger, Susanne (2007): ULME III. Anlage und Durchführung der Untersuchung. In: Lehmann, Rainer/Seeber, Susan (Hrsg.): ULME III. Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen. Hamburg. S. 21–40.

Lehmann, Rainer/Ivanov, Stanislav/Hunger, Susanne/Gänsfuß, Rüdiger (2005): ULME I. Untersuchung der Leistungen, Motivation und Einstellungen zu Beginn der beruflichen Ausbildung.

- Lehmann, Rainer/Seeber, Susan (Hrsg.) (2007): ULME III. Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen. Hamburg.
- Lehmann, Rainer/Seeber, Susan/Hunger, Susanne (2006): ULME II. Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der teilqualifizierenden Berufsfachschulen. Hamburg.
- Lehmann, Rainer/Seeber, Susan/Hunger, Susanne (2007): ULME III. Ziele der Untersuchung. In: Lehmann, Rainer/Seeber, Susan (Hrsg.): ULME III. Untersuchung von Leistungen, Motivation und Einstellungen der Schülerinnen und Schüler in den Abschlussklassen der Berufsschulen. Hamburg. S. 15–20.
- Lenhard, Wolfgang/Schneider, Wolfgang (Hrsg.) (2009): Diagnostik und Förderung des Leseverständnisses. Göttingen: Hogrefe.
- Leopold, Claudia (2009): Lernstrategien und Textverstehen. Spontaner Einsatz und Förderung von Lernstrategien. Münster: Waxmann.
- Loftus, Geoffrey (1972): Eye fixations and recognition memory for pictures. In: Cognitive Psychology 3 1972. S. 525–551.
- Loftus, Geoffrey/Kallman, Howard (1979): Encoding and use of detail information in picture recognition. In: Journal of Experimental Psychology: Human Learning and Memory 5 1979. S. 1979–211.
- McClelland, James/Rumelhart, David (1981): An Interactive Activation Model of Context Effects in Letter Perception. Part 1. An Account of Basic Findings. In: Psychological Review 88 1981. S. 375–407.
- McKoon, Gail/Ratcliff, Roger (1992): Inference during reading. In: Psychological Review 99 1992. S. 440–466.
- Merten, Klaus (1983): Inhaltsanalyse. Einführung in Theorie, Methode und Praxis. Opladen: Westdeutscher Verlag.
- Meyer, Bonnie (1975): The Organization of Prose and Its Effects on Memory. Amsterdam: North Holland.
- Michael, Berthold (1983): Darbieten und Veranschaulichen. Möglichkeiten und Grenzen von Darbietung und Anschauung im Unterricht. Bad Heilbrunn: Klinkhardt.
- Mikulecky, Larry (1988): Literacy for the workplace: The work education bridge. Bloomington: Indiana University.
- Mikulecky, Larry/Drew, Rad (1991): Basic Literacy Skills in the workplace. In: Kamil, Michael/Barr, Rebecca/Mosenthal, Peter/Pearson, David (Hrsg.): Handbook of Reading Research Vol. II 3. Aufl. Mahwah, NJ, London: Erlbaum. S. 669–689.
- Mikulecky, Larry/Ehlinger, Jeanne/Meenan, Avis (1987): Training for job literacy demands: What research applies to practice. University Park: Institute for the Study of Adult Literacy.
- Ministerium für Arbeit, Gesundheit und Soziales des Landes Nordrhein-Westfalen (2006): Praktische Altenpflegeausbildung in NRW. Lernort Praxis und Rahmenlehrplan.

Ministerium für Schule und Weiterbildung (2010): Lehrplan für das Berufskolleg in Nordrhein-Westfalen Medizinische Fachangestellte/Medizinischer Fachangestellter. Fachklassen des Dualen Systems der Berufsausbildung.

Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen (2003): Lehrplan zur Erprobung für den Ausbildungsberuf Industriekauffrau/ Industriekaufmann.

Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen (2004a): Lehrplan zur Erprobung für das Berufskolleg in Nordrhein-Westfalen Kauffrau für Spedition und Logistikdienstleistung/Kaufmann für Spedition und Logistikdienstleistung. Fachklassen des dualen Systems der Berufsausbildung.

Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen (2004b): Lehrplan zur Erprobung für den Ausbildungsberuf Kraftfahrzeugmechatronikerin/Kraftfahrzeugmechatroniker.

Ministerium für Schule, Jugend und Kinder des Landes Nordrhein-Westfalen (2004c): Lehrplan zur Erprobung für die industriellen Elektroberufe.

Möhn, Dieter (1991): Instruktionstexte. Ein Problemfall bei der Textidentifikation. In: Brinker, Klaus (Hrsg.): Aspekte der Textlinguistik. Hildesheim: Georg Olms Verlag. S. 183–212.

Moosbrugger, Helfried (2012a): Item-Response-Theorie (IRT). In: Moosbrugger, Helfried/Kelava, Augustin (Hrsg.): Testtheorie und Fragebogenkonstruktion 2. Aufl. Berlin Heidelberg: Springer Verlag. S. 227–274.

Moosbrugger, Helfried (2012b): Klassische Testtheorie. In: Moosbrugger, Helfried/Kelava, Augustin (Hrsg.): Testtheorie und Fragebogenkonstruktion 2. Aufl. Berlin Heidelberg: Springer Verlag. S. 103–117.

Morton, John (1969): The interaction of information in word recognition. In: Psychological Review 76 1969. S. 165–178.

Muckenhaupt, Manfred (1986): Text und Bild. Grundfragen der Beschreibung von Text-Bild-Kommunikationen aus sprachwissenschaftlicher Sicht. Tübingen: Gunter Narr Verlag.

Mugerauer, Roland (2012): Kompetenzen als Bildung? Die neuere Kompetenzenorientierung im Deutschen Schulwesen - eine skeptische Stellungnahme. Marburg: tectum.

Naceur, Abdelmajid (2001): Interesse und Textverstehen. Eine Untersuchung zum Einfluss des thematischen Interesses und kognitiver Faktoren auf die Repräsentation und das langfristige Behalten von Textinformationen. Bielefeld.

Naumann, Johannes/Artelt, Cordula/Schneider, Wolfgang/Stanat, Petra (2010): Lesekompetenz von PISA 2000 bis PISA 2009. In: Klieme, Eckhard/Artelt, Cordula/Hartig, Johannes/Jude, Nina/Köller, Olaf/Prenzel, Manfred/Schneider, Wolfgang/Stanat, Petra (Hrsg.): PISA 2009. Bilanz nach einem Jahrzehnt. Münster: Waxmann.

Nelson, Douglas/Brooks, David (1973): Functional independence of pictures and their verbal memory codes. In: Journal of Experimental Psychology 1 (98) 1973. S. 44–48.

Nelson, Douglas/Reed, Valerie/Walling, John (1976): Pictorial superiority effect. In: Journal of Experimental Psychology: Human Learning and Memory 5 (2) 1976. S. 523–528.

- Niederhaus, Constanze (2011): Die Komplexität von Fachtexten verschiedener Berufsfelder. In: Granato, Mona/Münk, Dieter/Weiß, Reinhold (Hrsg.): Migration als Chance. Bonn. S. 209–224.
- Niegemann, Helmut./Hessel, Silvia/Hochscheid-Mauel, Dirk/Aslanski, Kristina/Deimann, Markus/Kreuzberger, Gunther (2004): Kompendium E-Learning. Berlin: Springer Verlag.
- Nold, Günter/Willenberg, Heiner (2007): Lesefähigkeit. In: Beck, Bärbel (Hrsg.): Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch-Englisch-Schülerleistungen-International). Weinheim, Basel: Beltz. S. 23–41.
- Norwig, Kerstin, et al. (2013). Förderung der Lesekompetenz mittels Reciprocal Teaching - auch in der beruflichen Bildung ein Erfolg?. In: Zeitschrift für Berufs- und Wirtschaftspädagogik 1 (109) 2013. S. 67-93.
- Notter, Phillip/Arnold, Claudia/von Erlach, Emanuel/Hertig, Philippe (2006): Lesen und Rechnen im Alltag. Grundkompetenzen von Erwachsenen in der Schweiz. Nationaler Bericht zur Erhebung ALL. Neuchâtel.
- Notter, Phillip/Bonerad, Eva-Marie/Stoll, Francois (Hrsg.) (1999): Lesen - eine Selbstverständlichkeit? Schweizer Bericht zum "International Adult Literacy Survey". Zürich: Verlag Ruegger.
- OECD (Hrsg.) (1995): Literacy, Economy and Society. Results of the first International Adult Literacy Survey.
- OECD (2001): Knowledge and Skills for life. First Results from the OECD Programme for international student assessment (PISA) 2000: OECD Publishing.
- OECD (2005): Learning a living. First results of the adult literacy and life skills survey. Ottawa, Paris: OECD Publishing.
- OECD (2009): PIAAC Literacy. A conceptual framework.
- OECD (2012): Literacy, Numeracy and Problem Solving in Technology-Rich Environments. Framework for the OECD Survey of Adult Skills. Paris: OECD Publishing.
- OECD/Statistics Canada (Hrsg.) (1997): Literacy Skills for the Knowledge Society. Further Results from the International Adult Literacy Study. Paris, Ottawa.
- Oksaar, Els (1998): Fachsprachen und öffentliches Leben. Kommunikation in der arbeitsteiligen Gesellschaft. In: Hoffmann, Lothar/Kalverkämper, Hartwig/Wiegand, Herbert (Hrsg.): Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft. Berlin, New York: Walter de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft). S. 801–809.
- Oldenburg, Hermann (1992): Angewandte Fachtextlinguistik. „Conclusions“ und Zusammenfassungen. Tübingen.
- Olson, David (1994): The world on Paper. The conceptual and cognitive implications of writing and reading. Cambridge: Cambridge University Press.
- Paivio, Allan (1969): Mental imagery in associative learning and memory. In: Psychological Review 3 1969. S. 241–263.
- Paivio, Allan (1971): Imagery and Verbal Processes. New York: Holt, Rinehart and Winston.

- Paivio, Allan (1986): Mental representations. A dual coding approach. Oxford: Oxford University Press.
- Paivio, Allan (1990): Mental Representations. A dual Coding Approach. New York: Oxford University Press.
- Palincsar, A./Brown, A. (1984): Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. In: Cognition and Instruction 1 (2) 1984. S. 117–175.
- Pätzold, Günter (2010): Sprache - das kulturelle Kapital für eine Bildungs- und Berufskarriere. Sprachkompetenz und Berufsbildungskarriere. In: Zeitschrift für Berufs- und Wirtschaftspädagogik 2 2010. S. 161–171.
- Peirce, Charles (1906): Prolegomena for an Apology to Pragmaticism. In: The Monist 4 (16) 1906. S. 492–546.
- Peirce, Charles (1977): Semiotic and Significs. The Correspondence between Charles S. Peirce and Victoria Lady Welby. Hrsg. v. Charles S. Hardwick. Bloomington, London: Indiana University Press.
- Peirce, Charles (1986-1994): Semiotische Schriften. Vol. I-III. Hrsg. und übers. von Christian Kloesel und Helmut Pape. Frankfurt a.M.: Suhrkamp.
- Peirce, Charles (1993): Phänomen und Logik der Zeichen. 1. Aufl. 1983. Frankfurt a.M.: Suhrkamp.
- Pelka, Roland (1982): Sprachliche Aspekte von Bedienungsanleitungen technischer Geräte und Maschinen. In: Grosse, Siegfried/Mentrup, Wolfgang (Hrsg.): Anweisungstexte. Tübingen: Gunter Narr Verlag (Forschungsbericht des Instituts für deutsche Sprache). S. 74–103.
- Perrig, Walter/Kintsch, Walter (1985): Propositional and Situational Representations of Text. In: Journal of Memory and Language 5 (24) 1985. S. 503–518.
- Petsch, Cordula/Ziegler, Birgit/Gschwendtner, Tobias/Abele, Stephan/Nickolaus, Reinhold (2008): Lesekompetenzförderung in der beruflichen Bildung. In: Berufs- und Wirtschaftspädagogik - online 14 2008. S. 1–19.
- Pörksen, Uwe (1997): Weltmarkt der Bilder. Eine Philosophie der Visiotype. Stuttgart.
- Power, Sarah (1983): The Politics of literacy. In: Bailey, R./Fosheim, R. (Hrsg.): Literacy for Life. The Demand for Reading and Writing. New York: Modern Language Association of America. S. 21–29.
- Rammstedt, Beatrice (2013): PIAAC 2012. Die wichtigsten Ergebnisse im Überblick. In: Rammstedt, Beatrice (Hrsg.): Grundlegende Kompetenzen Erwachsener im internationalen Vergleich. Ergebnisse von PIAAC 2012. Münster, New York, München, Berlin: Waxmann. S. 11–20.
- Rammstedt, Beatrice/Zabal, Anouk (2013): Das Programme for the International Assessment of Adult Competencies (PIAAC). In: Rammstedt, Beatrice (Hrsg.): Grundlegende Kompetenzen Erwachsener im internationalen Vergleich. Ergebnisse von PIAAC 2012. Münster, New York, München, Berlin: Waxmann. S. 21–29.
- Raupp, Juliana/Vogelsang, Jens (2009): Medienresonanzanalyse. Eine Einführung in die Theorie und Praxis. Wiesbaden: Springer Verlag.

- Rayner, Keith/Pollatsek, Alexander (1989): The psychology of reading. Englewood Cliffs, NJ: Prentice Hall.
- Reinberg, Alexander/Hummel, Markus (2005): Höhere Bildung schützt auch in der Krise vor Arbeitslosigkeit. Vertrauter Befund. In: IAB Kurzbericht 9 2005. S. 1–6.
- Reinberg, Alexander/Hummel, Markus (2007): Der Trend bleibt – Geringqualifizierte sind häufiger arbeitslos. Schwierige Fortschreibung. In: IAB Kurzbericht 18 2007. S. 1–6.
- Rexing, Volker/Keimes, Christina/Ziegler, Birgit (2013): Lesekompetenz von BerufsschülerInnen. Befunde und Konsequenzen. In: Efing, Christian (Hrsg.): Ausbildungsvorbereitung im Deutschunterricht der Sekundarstufe I. Die sprachlich-kommunikativen Facetten von "Ausbildungsfähigkeit". Frankfurt am Main: PETER LANG (Wissen - Kompetenz - Text). S. 41–63.
- Richter, Tobias/Christmann, Ursula (2002): Lesekompetenz. Prozessebenen und interindividuelle Unterschiede. In: Groeben, Norbert/Hurrelmann, Bettina (Hrsg.): Lesekompetenz. Bedingungen, Dimensionen, Funktionen. Weinheim: Juventa. S. 25–58.
- Rickheit, Gert (1995): Verstehen und Verständlichkeit von Sprache. In: Spillner, Bernd (Hrsg.): Sprache: Verstehen und Verständlichkeit. Kongressbeiträge zur 25. Jahrestagung der Gesellschaft für angewandte Linguistik GAL e.V. Frankfurt am Main: PETER LANG (Forum Angewandte Linguistik). S. 15–30.
- Rickheit, Gert/Strohner, Hans (1993): Grundlagen der kognitiven Sprachverarbeitung. Modelle, Methoden, Ergebnisse. Tübingen: Francke.
- Roelcke, Thorsten (2005): Fachsprachen. Berlin: Schmidt Erich Verlag.
- Rössler, Patrick (2005): Inhaltsanalyse. Konstanz:: UTB.
- Rost, Jürgen (2004): Lehrbuch Testtheorie - Testkonstruktion. 2. Aufl. Bern, Göttingen, Toronto, Seattle: Hans Huber Verlag.
- Roth, Heinrich (1971): Pädagogische Anthropologie. Band 2. Entwicklung und Erziehung. Grundlagen einer Entwicklungspädagogik. Hannover: Schroedel.
- Rumelhart, D. (1994): Toward an interactive model of reading. In: Ruddell, R./Ruddell, M./Singer, H. (Hrsg.): Theoretical Models and Processes of Reading 4. Aufl. Newark: International Reading Association. S. 864–894.
- Rumelhart, David (1980): Schemata. The building blocks of cognition. In: Spiro, R.I./Bruce, B./Brewer, W. (Hrsg.): Theoretical Issues in reading comprehension. Perspectives from cognitive psychology, linguistics, artificial intelligence, and education. Hillsdale, N.J: Erlbaum. S. 33–58.
- Sadoski, Mark/Paivio, Allan (2001): Imagery and Text. A Dual Coding Theory of Reading and Writing. Mahwah: Lawrence Erlbaum.
- Sadoski, Mark/Paivio, Allan (2004): A Dual Coding Theoretical Model of Reading. In: Ruddell/Unrau (Hrsg.): Theoretical models and processes of reading 5. Aufl. Newark. S. 1329–1362.
- Sadoski, Mark/Paivio, Allan/Goetz, Ernest (1991): A critique of schema theory in reading and a dual coding alternative. In: Reading Research Quarterly 4 (26) 1991. S. 463–484.

Saile, Günter (1982): Wie montiert man einen Fleischwolf? Linguistische Analyse einer Anleitung. In: Grosse, Siegfried/Mentrup, Wolfgang (Hrsg.): Anweisungstexte. Tübingen: Gunter Narr Verlag (Forschungsbericht des Instituts für deutsche Sprache). S. 134–158.

Sauer, Christoph (1995): Ein Minimalmodell zur Verständlichkeitsanalyse und -optimierung. In: Spillner, Bernd (Hrsg.): Sprache: Verstehen und Verständlichkeit. Kongressbeiträge zur 25. Jahrestagung der Gesellschaft für angewandte Linguistik GAL e.V. Frankfurt am Main: PETER LANG (Forum Angewandte Linguistik). S. 149–171.

Sauer, Christoph (1999): Die Verständlichkeit von Texten, Visualisierungen und Bildschirmen. Untersuchungen zur Leseaufgabenunterstützung. In: Jakobs, Eva-Maria/Knorr, Dagmar/Pogner, Karl-Heinz (Hrsg.): Textproduktion. HyperText, Text, KonText. Frankfurt am Main, Berlin, Bern, New York, Paris, Wien: PETER LANG (Textproduktion und Medium). S. 93–109.

Schank, Roger (1982): Reading and understanding. Teaching from the Perspective of Artificial Intelligence. Hillsdale, N.J: Lawrence Erlbaum.

Schiefele, Ulrich (1996): Motivation und Lernen mit Texten. Göttingen: Hogrefe.

Schlagmüller, Matthias/Schneider, Wolfgang (2007a): Würzburger Lesestrategie-Wissenstest für die Klassen 7-12. Ein Verfahren zur Erfassung metakognitiver Kompetenzen bei der Verarbeitung von Texten. Göttingen: Hogrefe Verlag.

Schlagmüller, Matthias/Schneider, Wolfgang (2007b): Würzburger Lesestrategie-Wissenstest für die Klassen 7-12 (WLST 7-12). Manual. Göttingen: Hogrefe.

Schmalhofer, Franz/Glavanov, Doris (1986): Three Components of understanding a programmer's manual. Verbatim, Positional, and Situational Representations. In: Journal of Memory and Language 25 1986. S. 279–294.

Schmidt, Julie/McLaughlin, John/Lehten, Patricia (1989): Novice strategies for understanding paintings. In: Applied Cognitive Psychology 3 1989. S. 65–72.

Schneider, Wolfgang (2009): Diagnostik basaler Lesekompetenzen. In: Lenhard, Wolfgang/Schneider, Wolfgang (Hrsg.): Diagnostik und Förderung des Leseverständnisses. Göttingen: Hogrefe. S. 45–64.

Schnotz, Wolfgang (1988): Textverstehen als Aufbau mentaler Modelle. In: Mandl, Heinz/Spada, Hans (Hrsg.): Wissenspsychologie. München, Weinheim: Psychologie Verlags Union. S. 299–330.

Schnotz, Wolfgang (1994): Aufbau von Wissensstrukturen. Untersuchungen zur Kohärenzbildung beim Wissenserwerb mit Texten. Weinheim: Beltz.

Schnotz, Wolfgang (1997): Wissenserwerb mit Diagrammen und Texten. In: Issing, Ludwig/Klimsa, Paul (Hrsg.): Information und Lernen mit Multimedia und Internet. Lehrbuch für Studium und Praxis 2. Aufl. Weinheim: Beltz. S. 85–105.

Schnotz, Wolfgang (2003): Informationsintegration mit Sprache und Bild. In: Rickheit, Gert/Herrmann, Theo/Deutsch, Werner (Hrsg.): Psycholinguistik. Ein internationales Handbuch. Berlin: Walter de Gruyter. S. 577–587.

Schnotz, Wolfgang (2005): An integrated Modell of Text and Picture Comprehension. In: Mayer, Richard (Hrsg.): The Cambridge Handbook of multimedia Learning. Cambridge: University Press. S. 49–69.

- Schnotz, Wolfgang (2006a): Textverständnis. In: Rost, Detlef (Hrsg.): Handwörterbuch Pädagogische Psychologie 3. Aufl. Weinheim: Beltz. S. 769–778.
- Schnotz, Wolfgang (2006b): Was geschieht im Kopf des Lesers? Mentale Konstruktionsprozesse beim Textverstehen aus der Sicht der Psychologie und der kognitiven Linguistik. In: Blühdorn, Hardarik/Breindl, Eva/Waßer, Ulrich (Hrsg.): Text - Verstehen. Grammatik und darüber hinaus. Berlin, New York: Walter de Gruyter. S. 224–238.
- Schnotz, Wolfgang (2010): Wissenserwerb mit Texten, Bildern und Diagrammen. In: Issing, Ludwig/Klimsa, Paul (Hrsg.): Information und Lernen mit Multimedia und Internet. Lehrbuch für Studium und Praxis 3. Aufl. Weinheim: Beltz. S. 65–81.
- Schnotz, Wolfgang/Baadte, C./Müller, Astrid/Rasch, R. (2011): Kreatives Denken und Problemlösen mit bildlichen und beschreibenden Repräsentationen. In: Sachs-Hombach, R./Totzke, R. (Hrsg.): Bilder-Sehen-Denken. Köln. S. 204–252.
- Schnotz, Wolfgang/Bannert, Maria (1999): Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. In: *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)* 3 (46) 1999. S. 217–236.
- Schnotz, Wolfgang/Bannert, Maria (2003): Construction and inference in learning from multiple representation. In: *Learning and Instruction* (13) 2003. S. 141–156.
- Schnotz, Wolfgang/Dutke, Stephan (2004): Kognitionspsychologische Grundlagen der Lesekompetenz. Mehrebenenverarbeitung anhand multipler Informationsquellen. In: Schiefele, Ulrich/Artelt, Cordula/Schneider, Wolfgang/Stanat, Petra (Hrsg.): Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000. Wiesbaden: VS Verlag für Sozialwissenschaften. S. 61–100.
- Schnotz, Wolfgang/Horz, Holger/McEvang, Nele/Schroeder, Sascha/Ulrich, Mark/Baumert, Jürgen/Hackfeld, Axinja/Richter, Tobias (2010): Das BITE-Projekt. Integrative Verarbeitung von Bildern und Texten in der Sekundarstufe I. In: Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. In: *Zeitschrift für Pädagogik* (56) 2010. S. 141–153.
- Schwarz, Maria/Flammer, August (1979): Erstinformationen einer Geschichte. Ihr Behalten und ihre Wirkung auf das Behalten der nachfolgenden Informationen. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 4 1979. S. 347–358.
- Seidenberg, Mark/Waters, Gloria/Barnes, Marcia/Tanehaus, Michael (1984): When does Irregular Spelling or Pronunciation Influence Word Recognition? In: *Journal of Verbal Learning and Verbal Behavior* 23 1984. S. 383–404.
- Simpson, Greg/Krueger, Merilee (1991): Selective Access of Homograph Meanings in Sentence Context. In: *Journal of Memory and Language* 30 1991. S. 627–643.
- Sloane, Peter/Dilger, Bernadette (2005): The Competence Clash. Dilemmata bei der Übertragung des 'Konzepts der nationalen Bildungsstandards' auf die berufliche Bildung. In: *Berufs- und Wirtschaftspädagogik online* 8 2005. S. 1–32.
- Smith, Cecil (2000): The real-world reading practices of adults. In: *Journal of Literacy Research* 1 (32) 2000. S. 25–52.

Smith, Frank (2004): Understanding reading. A Psycholinguistic Analysis of Reading and Learning to Read. 6. Aufl. Mahwah: Lawrence Erlbaum.

Solfjeld, Kare (2010): Die sprachliche Realisierung von Handlungsschritten in Gebrauchsanweisungen. Vergleich Deutsch - Norwegisch. In: Kvam, Sigmund/Knutsen, Karen Patrick/Langemeyer, Peter (Hrsg.): Textsorten und kulturelle Kompetenz. Interdisziplinäre Beiträge zur Textwissenschaft. Münster: Waxmann. S. 203–225.

Souvignier, Elmar/Trenk-Hinterberger, I./Adam-Schwebe, Stefanie/Gold, Andreas (2009): Frankfurter Leseverständnistest für 5. und 6. Klassen. Göttingen: Hogrefe Verlag.

Spoden, Christian/Frey, Andreas/Bernhardt, Raphael/Ziegler, Birgit/Seeber, Susan (2014/2015): Differenzielle Domänen- und Itemeffekte zwischen Ausbildungsberufen bei der Erfassung allgemeiner schulischer Kompetenzen von Berufsschülerinnen und Berufsschülern. (In Vorbereitung). In: Zeitschrift für Berufs- und Wirtschaftspädagogik (Sonderheft) 2014/2015.

Stalder, Barbara (2011): Berufsausbildung mit geringen Lesekompetenzen? Ergebnisse der Schweizer PISA-Folgestudie TREE. In: Fischer, Martin/Becker, Matthias/Spöttl, Georg (Hrsg.): Kompetenzdiagnostik in der beruflichen Bildung - Probleme und Perspektiven. Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien: PETER LANG (Berufliche Bildung in Forschung, Schule und Arbeitswelt). S. 190–203.

Standing, Lionel (1973): Learning 10000 pictures. In: Quarterly Journal of experimental Psychology (25) 1973. S. 207–222.

Standing, Lionel/Conezio, J./Haber, R. (1970): Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. In: Psychonomic Science (19) 1970. S. 73–74.

Statistics Canada (1996): Introduction. In: Statistics Canada (Hrsg.): Reading the Future. A Portrait of Literacy in Canada. Ottawa.

Statistisches Bundesamt (Hrsg.) (2006): Verdienststrukturerhebung. Vollzeitbeschäftigte Arbeitnehmer. Pressemitteilung.

Statistisches Bundesamt (Hrsg.) (2. Quartal 2013): Verdienste und Arbeitskosten. Arbeitnehmerverdienste.

Stein, Peter (2010): Schriftkultur. Eine Geschichte des Schreibens und Lesens. 2. Aufl. Darmstadt: Wissenschaftliche Buchgesellschaft.

Steyer, Rolf/Eid, Michael (1993): Messen und Testen. Berlin: Springer Verlag.

Sticht, Thomas (1975): Reading for Working. A Functional Literacy Anthology. Alexandria: Human Resources Research Organization.

Sticht, Thomas (1977): Comprehending Reading at Work. In: Just, Marcel/Carpenter, Patricia (Hrsg.): Cognitive processes in comprehension. Hillsdale, NJ: Lawrence Erlbaum. S. 221–246.

Sticht, Thomas/Mikulecky, Larry (1984): Job-Related Basic Skills. Cases and Conclusions. Columbus: National Center Publications.

Stöckl, Hartmut (2004): Die Sprache im Bild - das Bild in der Sprache. Zur Verknüpfung von Sprache und Bild im massenmedialen Text. Konzepte - Theorien - Analysemethoden. Berlin: De Gruyter.

- Stoll, Francois (1999): Zur Bedeutung des Lesens und Schreibens. In: Notter, Phillip/Bonerad, Eva-Marie/Stoll, Francois (Hrsg.): Lesen - eine Selbstverständlichkeit? Schweizer Bericht zum "International Adult Literacy Survey". Zürich: Verlag Ruegger. S. 25–30.
- Straßner, Erich (2002): Text-Bild-Kommunikation Bild-Text-Kommunikation. Tübingen: Niemeyer.
- Taft, Marcus (1979): Lexical Access Via an Orthographic Code. The Basis Orthography Syllable Structure (BOSS). In: Journal of Verbal Learning and Verbal Behavior 18 1979. S. 21–38.
- Taft, Marcus (1986): Lexical Access Codes in Visual and Auditory Word Recognition. In: Language and Cognitive Processes 4 1986. S. 297–308.
- Taft, Marcus/Foster, Kenneth (1975): Lexical Storage and Retrieval of Prefixes Words. In: Journal of Verbal Learning and Verbal Behavior 14 1975. S. 638–647.
- Teigeler, Peter (1968): Verständlichkeit und Wirksamkeit von Sprache und Text. Stuttgart: Nadolski.
- Tröster, Monika (Hrsg.) (2002): Berufsorientierte Grundbildung. Konzepte und Praxishilfen. Bielefeld: Bertelsmann Verlag.
- Vonken, Matthias (2005): Handlung und Kompetenz. Theoretische Perspektiven für die Erwachsenen- und Berufspädagogik. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Vonken, Matthias (2011): Kritische Anmerkungen zum Kompetenzbegriff. In: Bethschneider, Monika/Höhns, Gabriela/Münchhausen Gesa (Hrsg.): Kompetenzorientierung in der beruflichen Bildung. Berichte zur beruflichen Bildung. Bonn: Bundesinstitut für Berufsbildung (Schriftenreihe des Bundesinstituts für Berufsbildung). S. 21–32.
- Wagner, Hildegard (1984): Die deutsche Verwaltungssprache der Gegenwart. Eine Untersuchung der sprachlichen Sonderform und ihrer Leistung. Düsseldorf: Schwann.
- Waldmann, Michael (1990): Schema und Gedächtnis. Das Zusammenwirken von Raum- und Ereignisschemata beim Gedächtnis für Alltagssituationen. Heidelberg: Asanger.
- Weidenmann, Bernd (1988): Psychische Prozesse beim Verstehen von Bildern. Bern: Huber.
- Weidenmann, Bernd (1994): Wissenserwerb mit Bildern. Instruktionale Bilder in Printmedien, Film / Video und Computerprogrammen. Bern.
- Weissenberg, Jens (2010): Sprachlich-kommunikative Handlungsfelder am Arbeitsplatz. Konzeptioneller Ansatz zur Entwicklung, Durchführung und Evaluation berufsbezogener Zweitsprachangebote. In: Deutsch als Zweitsprache 2 2010. S. 13–24.
- Wengel, Peter (2013): Sprachlich-kommunikative Anforderungen in der Berufsschule. In: Efinger, Christian (Hrsg.): Ausbildungsvorbereitung im Deutschunterricht der Sekundarstufe I. Die sprachlich-kommunikativen Facetten von "Ausbildungsfähigkeit". Frankfurt am Main: PETER LANG (Wissen - Kompetenz - Text). S. 146–170.
- West, Caroline/O'Rourke, Timothy/Holcomb, Phillip (1998): Event-related brain potentials and language comprehension. A cognitive neuroscience approach to the study of intellectual functioning. In: Soraci, S./McIlvane (Hrsg.): Perspectives on fundamental processes in intellectual functioning. Stamford: Ablex. S. 131–168.

- White, Sheida/Chen, Jing/Forsyth, Barbara (2010): Reading-Related Literacy Activities of American Adults: Time Spent, Task Types, and Cognitive Skills Used. In: *Journal of Literacy Research* (42) 2010. S. 276–307.
- Wiese, Johannes (1978): Visual Education - Visual Literacy. Theorien und Praxisansätze für visuelle Kommunikation in USA und Großbritannien. In: Allendorf, Otmar/Wiese, Johannes (Hrsg.): *Taschenbuch der Kommunikationspraxis. Beispiele einer Film- und Fernseherziehung*. Köln: interorga (Theorie und Praxis der Unterrichtstechnologie). S. 65–70.
- Willenberg, Heiner (2005): Ein handhabbares System, um Textschwierigkeiten einzuschätzen. Vorschläge für eine Textdatenbank von Sachtexten. In: Fix, Martin/Jost, Roland (Hrsg.): *Sachtexte im Deutschunterricht*. Baltmannsweiler: Schneider Verlag Hohengehren. S. 94–106.
- Winther, Esther (2010): *Kompetenzmessung in der beruflichen Bildung*. Bielefeld: Bertelsmann Verlag.
- Wolff, Dieter (1985): Textverständlichkeit und Textverstehen. Wie kann man den Schwierigkeitsgrad eines authentischen fremdsprachlichen Textes bestimmen? In: *Neusprachliche Mitteilungen aus Wissenschaft und Praxis* 38 1985. S. 211–221.
- Wright, Peter/Lickorish, Ann (1983): Proof-reading texts on screen and paper. In: *Behaviour & Information Technology* 3 (2) 1983. S. 227–235.
- Wu, Margaret/Adams, Raymond/Wilson, Mark/Haldane, Samuel (2007): *ACER Conquest (Version 2.0). Generalized item response modelling software*. Camberwel: Australian Council for Educational Research.
- Ziegler, Birgit/Balkenhol, Aileen/Keimes, Christina/Rexing, Volker (2012): Diagnostik „funktionaler Lesekompetenz“. In: *Berufs- und Wirtschaftspädagogik - online* 22 2012. S. 1–19.
- Ziegler, Birgit/Gschwendtner, Tobias (2010): Leseverstehen als Basiskompetenz. Entwicklung und Förderung im Kontext beruflicher Bildung. In: *Zeitschrift für Berufs- und Wirtschaftspädagogik* 4 2010. S. 534–555.
- Zieten, Werner (1988): Betriebsanleitungen wie der Kunde sie wünscht. In: Bungarten, Theo (Hrsg.): *Sprache und Information in Wirtschaft und Gesellschaft*. Todstedt: Attikon Verlag. S. 253–261.
- Zwaan, Rolf (1992): Aspects of literary comprehension. A cognitive approach.
- Zwaan, Rolf (2004): The immersed experiencer. Toward an embodied theory of language comprehension. In: Ross, Brian (Hrsg.): *The Psychology of Learning and Motivation*. New York. S. 35–62.
- Zwaan, Rolf/Madden, C. (2005): Embodied sentence comprehension. In: Pecher, D./Zwaan, Rolf (Hrsg.): *The Grounding of Cognition. The role of Perception and Action in Memory, Language, and Thinking*. Cambridge.
- Zwitserlood, Piene (1995): Prozesse und lexikalische Repräsentationen bei der visuellen Worterkennung. In: Spillner, Bernd (Hrsg.): *Sprache: Verstehen und Verständlichkeit. Kongressbeiträge zur 25. Jahrestagung der Gesellschaft für angewandte Linguistik GAL e.V. Frankfurt am Main: PETER LANG (Forum Angewandte Linguistik)*. S. 115–118.

7. Anhang

7.1. Ratingbogen

Darbietung

Anzahl der PowerPoint Folien: _____

Art und Anzahl der Informationsquellen

- Eine Quelle
- mehrere Quellen des gleichen Repräsentationsmodus
- mehrere Quellen verschiedener Repräsentationsmodi

Das Format des Itemstimulus ist...

- Deskriptional
- Depiktional
- Gemischt

Verarbeitungsprozesse

- Es müssen keine Inferenzen gebildet werden
- 1-2 Inferenzen erforderlich
- 3 Inferenzen erforderlich

Moduswechsel zur Beantwortung notwendig?

- Nein
- Ja

Komplexität des Mentalen Modells

- kein Modell erforderlich
- einfach
- mittel
- komplex

weitere Merkmale

Platzierung der relevanten Informationen

- betont/auffällig

- offen
- versteckt

Anzahl der irrelevanten/konkurrierenden Informationen im präsentierten

Material

- Wenige
- Einige
- Viele

Die Antworten auf das Item erfolgen

- Geschlossen mit 1 plausiblen Antwortalternative
- Geschlossen mit 2 plausiblen Antwortalternativen
- Geschlossen mit 3 plausiblen Antwortalternativen
- offen

Wissensvoraussetzungen

Besonderes Sprachwissen (Text-/Grafikstrukturwissen – bspw. wie etwas grafisch dargestellt wird, bestimmte Textmerkmale,...) notwendig?

- Nein
- Ja

Besonderes Weltwissen (Wissen über die Sachverhalte des Textinhalts, inhaltliches Vorwissen) notwendig?

- Nein
- Ja

Das Item ist...

- einfach
- mittel
- schwer

7.2. Ergänzende Tabellen und Abbildungen

Tabelle 42 – Leseanlässe im Rahmenlehrplan für die berufliche Grundbildung des Ausbildungsberufs Maurer/in (Auszug) (Keimes/Rexing 2011b, S. 4)

Lernfeld	explizite Leseanlässe im Lehrplan	implizite Leseanlässe im Lehrplan
Einrichten einer Baustelle	Lesen von Plänen zur Baustelleneinrichtung	Planen eines Baustelleneinrichtungsplans unter Beachtung der Arbeitsschutzvorschriften bzw. Vorschriften des Umweltschutzes
Erschließen und Gründen eines Bauwerkes		Planen der Herstellung von Baugruben und Gräben unter Berücksichtigung der Unfallverhütungsvorschriften
Mauern eines einschaligen Baukörpers		Durchführen von Mengen- und Materialermittlungen anhand von Tabellen
Herstellen eines Stahlbetonbauteiles		Bestimmen der Zusammensetzung des Betons anhand von Tabellen
Herstellen einer Holzkonstruktion		Treffen von Entscheidungen zum Holzschutz
Beschichten und Bekleiden eines Bauteiles		Auswählen von Beschichtungs-, Bekleidungs- und Belagsmaterialien

Tabelle 43 – Übersicht über relevantes Textmaterial (Experteninterviews) (Keimes 2014, S. 154)

Hierarchie- ebene Textmaterial	Auszubildende	Gesellen (Facharbeiter)	Personen mit Führungs-verantwor- tung
(Ausführungs-) Zeichnungen	✓	✓	✓
Bedienungsanleitungen von Geräten	✓	✓	✓
Produkt-, Ausführungs- und Verarbeitungshinweise	✓	✓	✓
Sicherheitshinweise/-vor- schriften	✓	✓	✓
Leistungsverzeichnisse	✓	✓	✓
Tabellenwerke	✓	✓	✓
Arbeitsaufträge/-anweisun- gen	-	✓	✓
Lieferscheine	-	✓	✓
Materialzettel	-	✓	✓
Checklisten	-	✓	✓
gesetzliche Vorschriften	-	✓	✓
Fachzeitschriften	-	✓	✓
Normen	-	✓	✓
Bodengutachten	-	✓	✓
Statiken	-	✓	✓
Tagesberichte	-	✓	✓
Personaleinsatzpläne	-	-	✓
E-Mails/Schriftverkehr	-	-	✓
Genehmigungen	-	-	✓
Kalkulationen	-	-	✓

Tabelle 44 – Textrezeptive Handlungsfelder (Keimes 2014, S. 164)

textrezeptive Handlungsfelder	Bezeichnung	Textmaterial
HF 1	Arbeitsplanung/-organisation	(Ausführungs-)Zeichnungen Personaleinsatzpläne Bauzeitenpläne
HF 2	Materialbeschaffung und -annahme	Lieferscheine Materialzettel
HF 3	Ausführung/Erstellung von Bauteilen	Leistungsverzeichnisse Produkt-, Ausführungs- und Verarbeitungshinweise
HF 4	Arbeit mit Maschinen und Elektrogeräten	Bedienungsanleitungen von Geräten
HF 5	Gewährleistung der Sicherheit	Sicherheitshinweise/-vorschriften gesetzliche Vorschriften
HF 6	Reaktion auf Bauablaufstörungen	Bodengutachten Statiken Normen Tabellenwerke
HF 7	Qualitätskontrolle und -sicherung	Checklisten
HF 8	Kontrolle der Wirtschaftlichkeit	Kalkulationen
HF 9	Kommunikation mit internen Akteuren	Arbeitsaufträge/-anweisungen
HF 10	Kommunikation mit externen Akteuren	E-Mails/Schriftverkehr Genehmigungen
HF 11	Lehr-/Lernprozesse im Kontext Aus- und Weiterbildung	Tagesberichte Fachzeitschriften

Tabelle 45 – Einordnung der Texte in das Modell funktionale Lesekompetenz

<div>Repräsentati- onsformate</div> <div>Anforder- ungsklassen</div>	deskriptional	depiktional	gemischte Formate
identifizieren	E-Mails/Schriftver- kehr (HF 10) Genehmigungen (HF 10)		Personaleinsatzpläne (HF 1) Arbeitsaufträge/-anweisun- gen (HF 9) Checklisten (HF 7) Tagesberichte (HF 11)
integrieren	gesetzliche Vor- schriften (HF 5)	Sicherheitshinweise/- vorschriften (HF 5)	Materialzettel (HF 2) Kalkulationen (HF 8) Fachzeitschriften (HF 11)
generieren		(Ausführungs-) Zeich- nungen (HF 1)	Lieferscheine (HF 2) Bedienungsanleitungen von Geräten (HF 4) Produkt-, Ausführungs- und Verarbeitungshinweise (HF 3) Leistungsverzeichnisse (HF 3) Normen (HF 6) Tabellenwerke (HF 6) Bodengutachten (HF 6) Statiken (HF 6)

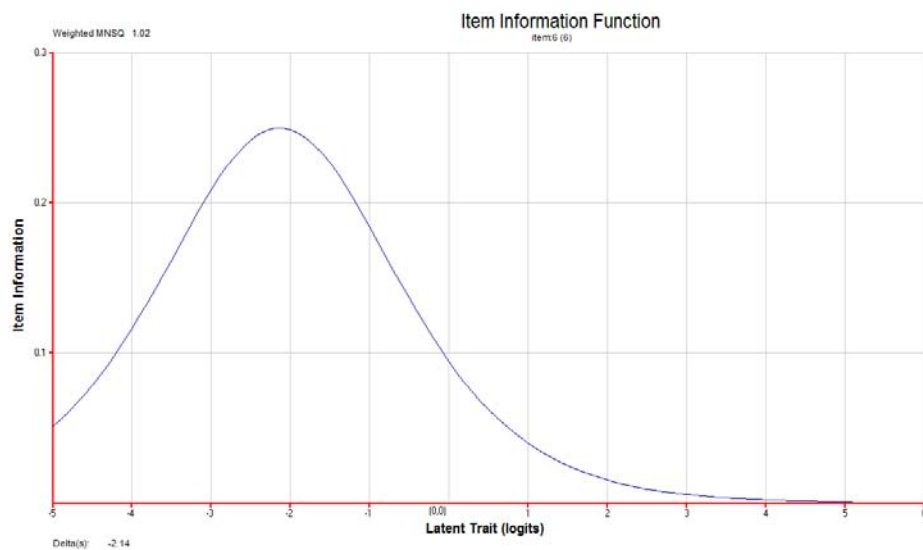


Abbildung 46 – IIF Item 6 (CBA)

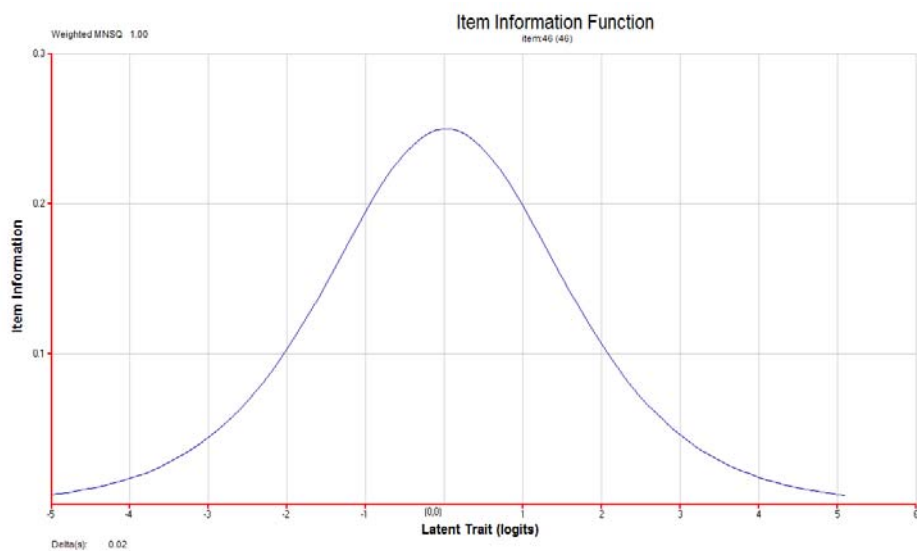


Abbildung 47 – IIF Item 49 (CBA)

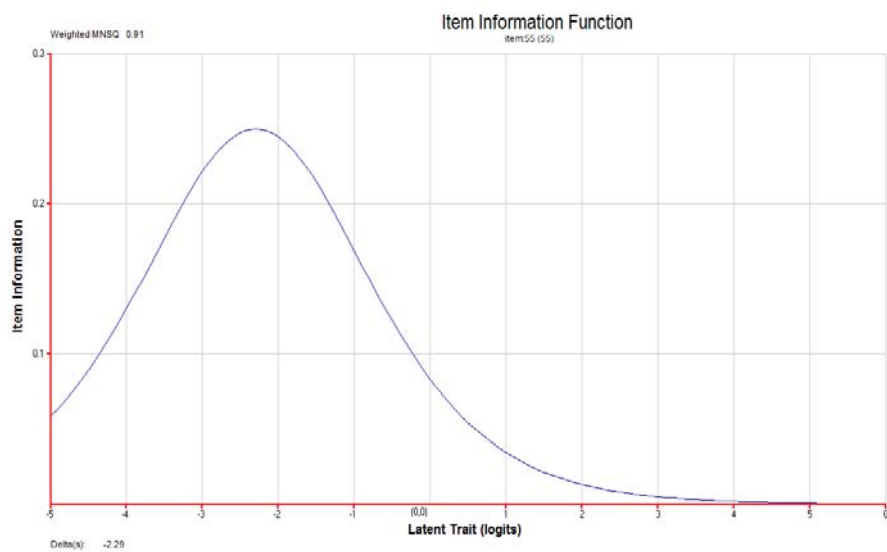


Abbildung 48 – IIF Item 59 (CBA)

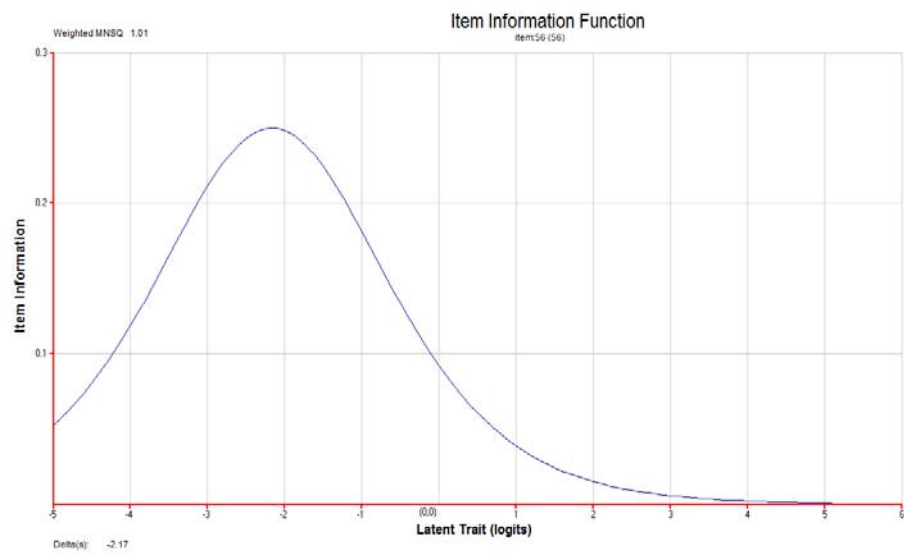


Abbildung 49 – IIF Item 60 (CBA)

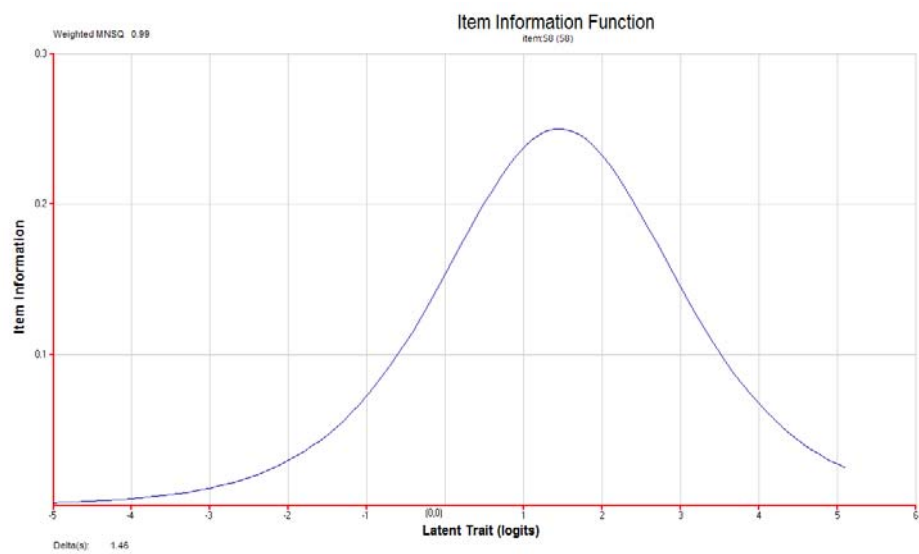


Abbildung 50 – IIF Item 62 (CBA)

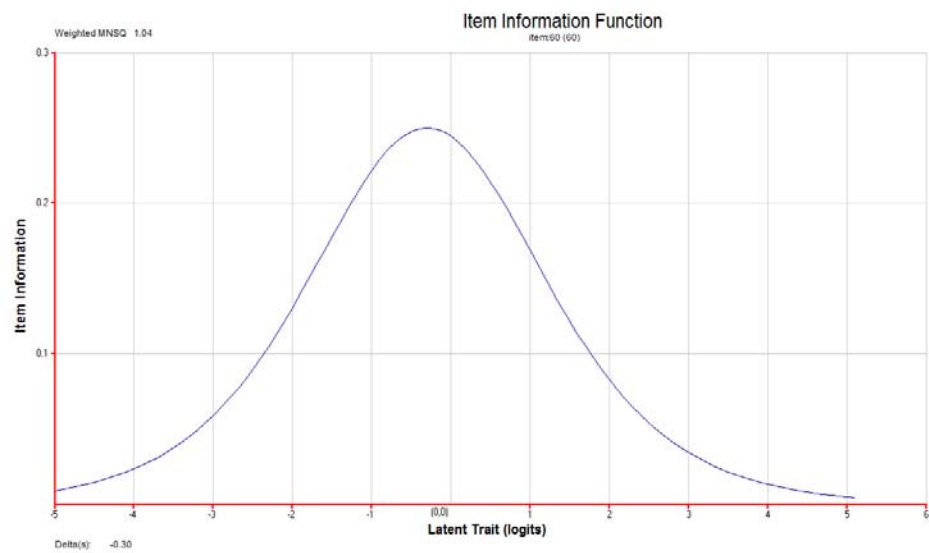


Abbildung 51 – IIF Item 64 (CBA)

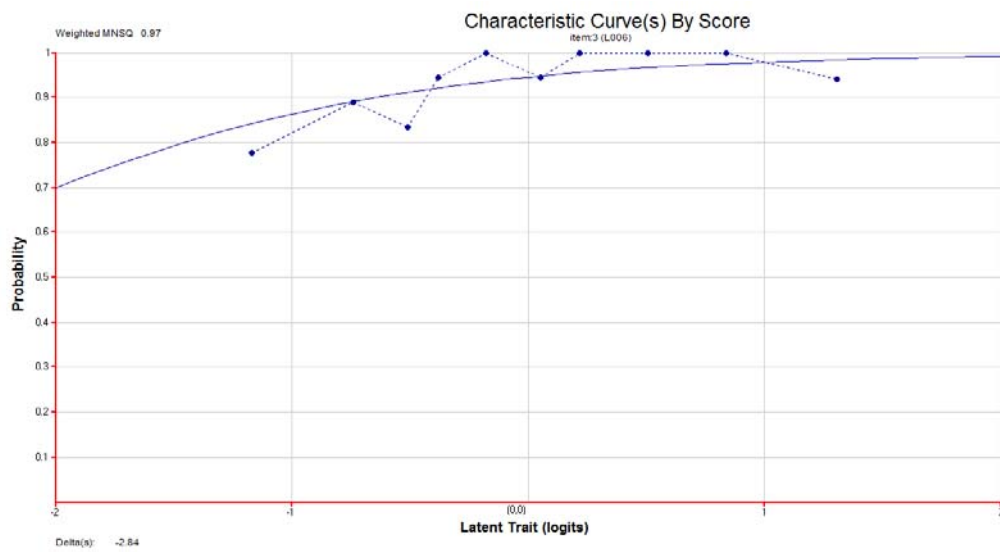


Abbildung 52 – ICC Item 6 (PP)

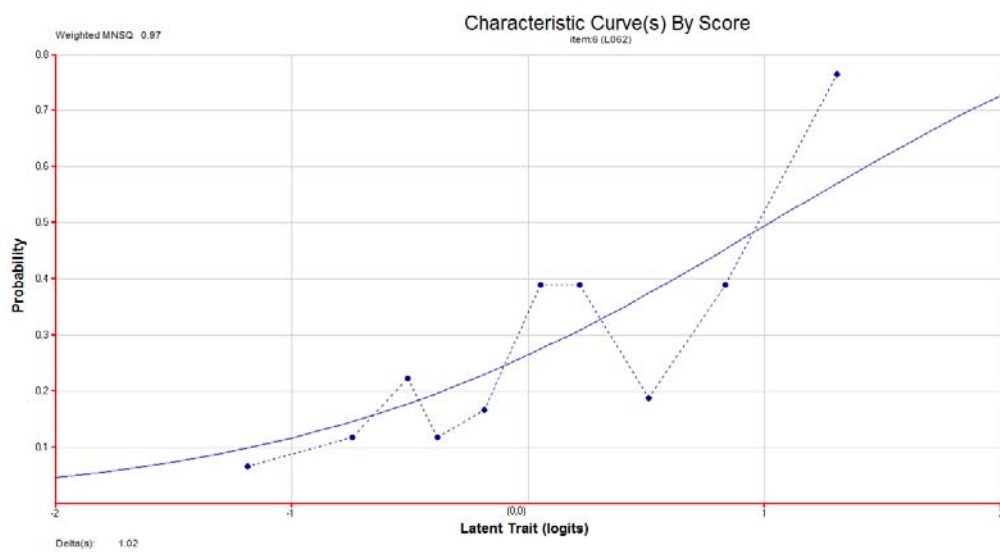


Abbildung 53 – ICC Item 62 (PP)

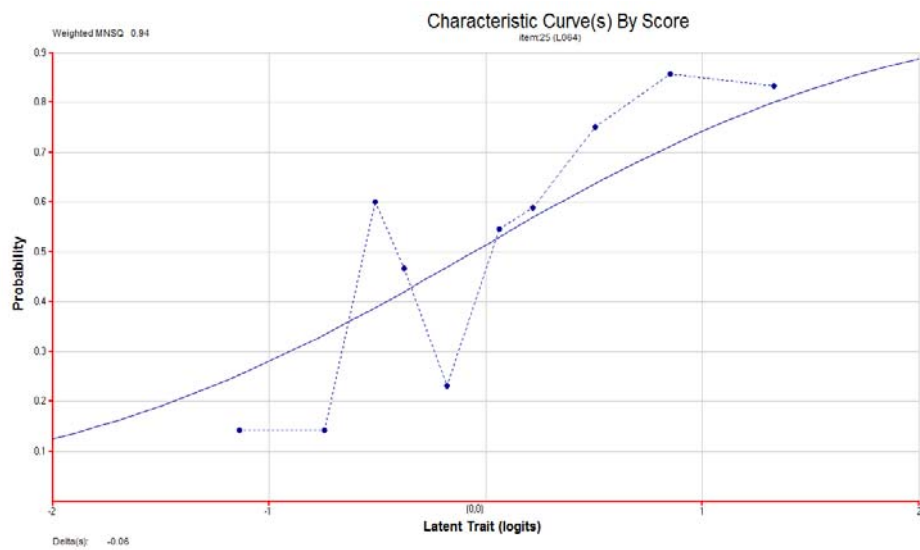


Abbildung 54 – ICC Item 64 (PP)

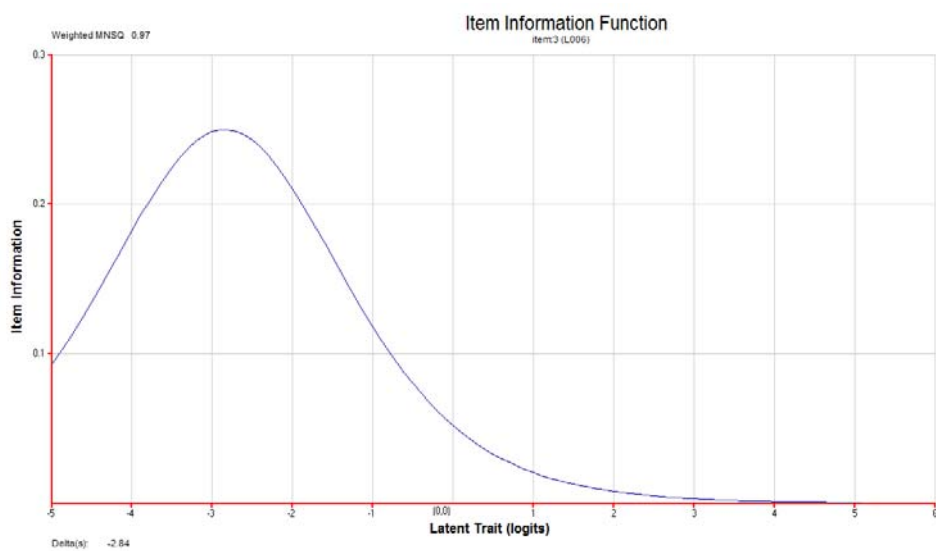


Abbildung 55 – IIF 6 (PP)

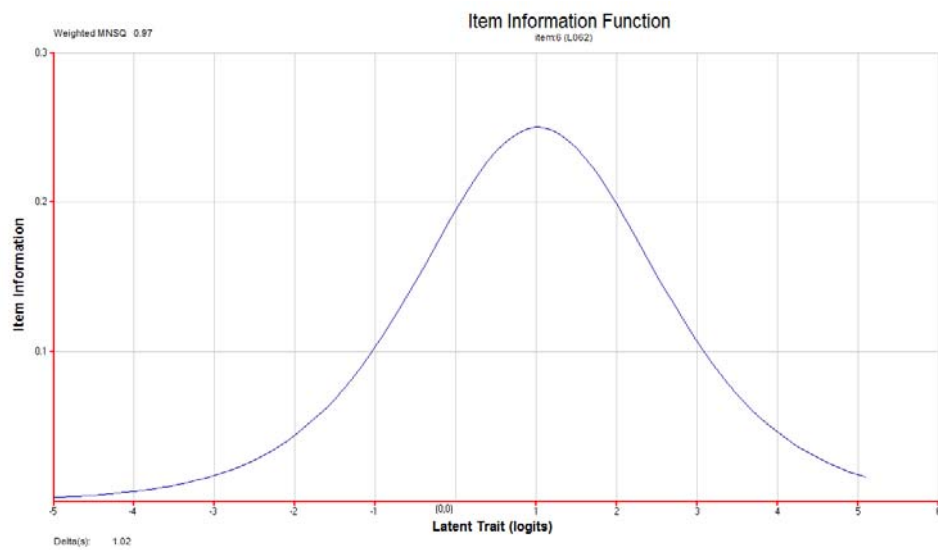


Abbildung 56 – IIF Item 62 (PP)

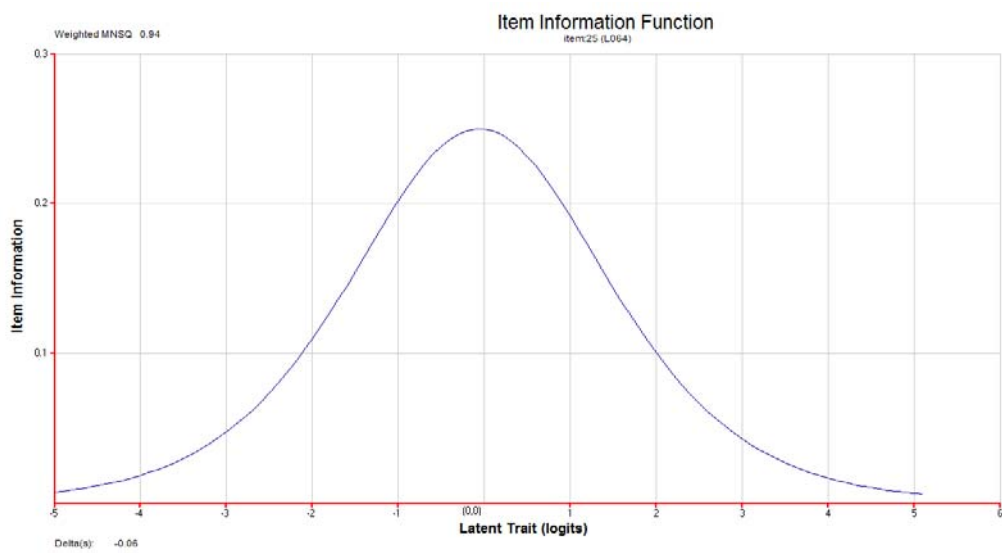


Abbildung 57 – IIF Item 64 (PP)

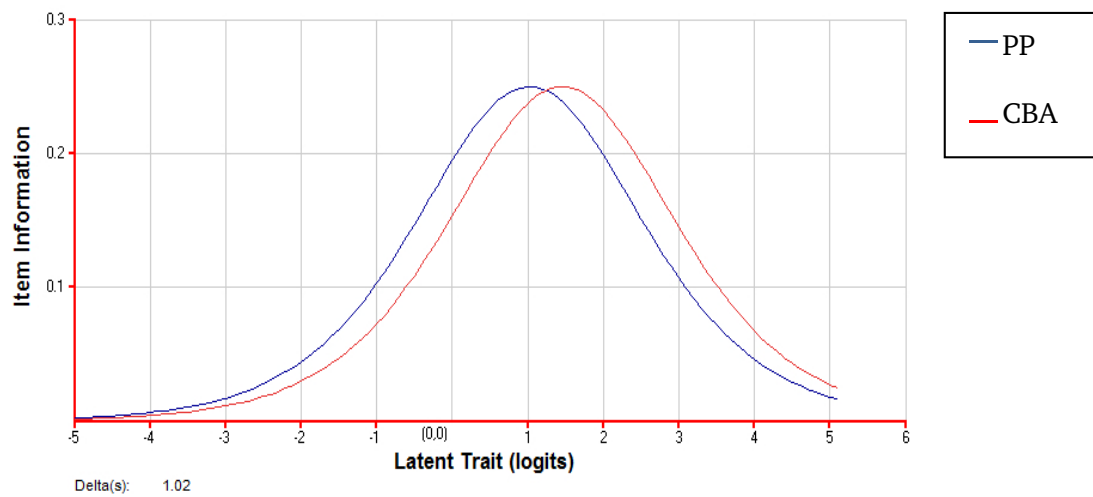


Abbildung 58 – IIFs Item 62 (CBA + PP)

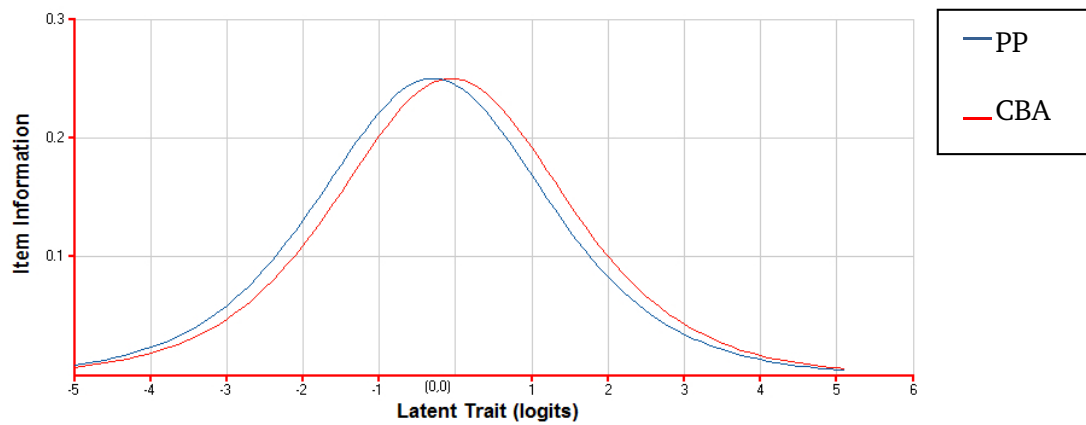


Abbildung 59 – IIFs Item 64 (CBA + PP)

Tabelle 46 – Itemkennwerte (CBA [blau] + PP)

ID	Estimate		Error		WMNSQ		t		Dicrimination	
	CBA	PP	CBA	PP	CBA	PP	CBA	PP	CBA	PP
1	.397	-3.792	.163	.462	1	1	.1	.10	.43	.51
4	-1.292	-2.919	.185	.322	.99	1.02	-.1	.20	.47	.33
6	-2.103	-2.843	.231	.31	1.03	.97	.3	.00	.36	.23
9	.864	.291	.169	.171	1	1	0	-.10	.46	.4
10	-1.161	-1.329	.182	.195	1	.98	0	-.20	.45	.26
11	-.644	1.023	.163	.187	.99	.97	-.2	-.30	.47	.33
17	-2.525	1.327	.261	.201	.96	1.05	-.2	.50	.42	.12
20	-1.177	-1.207	.18	.189	.98	.98	-.2	-.20	.4	.35
28	.457	1.899	.161	.225	.99	1	-.2	.00	.47	.23
29	-.685	-1.743	.166	.22	1.06	.97	1	-.20	.38	.34
36	1.331	1.050	.186	.186	1.02	1.02	.2	.30	.33	.35
37	-1.066	-2.304	.179	.256	.97	1	-.3	.10	.53	.52
38	.3	.633	.17	.177	1.01	1.01	.2	.20	.44	.37
39	-.732	1.198	.173	.192	.9	1.02	-1.5	.20	.65	.34
40	.645	.361	.167	.176	1.01	.99	.2	-.10	.44	.41
41	-.072	.786	.163	.179	1.04	.93	.8	-1.10	.42	.27
42	-.002	-.739	.171	.187	1.08	1	1.6	.00	.32	.38
45	-1.087	-.283	.182	.178	.99	1.05	-.1	.90	.43	.45
46	.374	-.204	.16	.183	1.02	1.01	.4	.20	.43	.33
48	-2.972	-.052	.31	.181	.98	1.03	0	.50	.29	.13
50	1.742	2.203	.203	.274	1	1	0	.00	.38	.31
52	1.66	.379	.202	.188	1	1.09	0	1.60	.37	.32
54	.425	-.056	.161	.19	1.01	.94	.2	-1.10	.48	.17
62	1.435	.477	.191	.199	.99	1	-.1	.10	.41	.4
63	1.823	-.174	.207	.196	1	1.03	.1	.50	.33	.32
64	-.297	-.634	.166	.204	1.05	.94	.9	-.80	.39	.47
66	-.276	1.091	.161	.22	1.01	1.1	.2	1.00	.44	.33
67	.184	-.856	.161	.21	.99	1.07	-.3	.90	.45	.35
69	-.628	-.750	.164	.211	.96	1.02	-.6	.30	.54	.32
70	.073	-2.396	.16	.331	1.01	.92	.2	-.30	.48	.33
72	-.748	-.141	.17	.204	1.06	.91	1	-1.60	.38	.23

Tabelle 47 – Ergebnisse CBA Kalibrierung für die im PP gelaufenen Items aufgeschlüsselt nach Anforderungsbereichen und Repräsentationsformaten

	Deskriptional	Hybrid	Depiktional
	Estimate <u>-1.10</u> Error <u>.17</u> WMNSQ <u>1.01</u> t <u>.17</u> DC <u>.43</u>	Estimate <u>-1.16</u> Error <u>.18</u> WMNSQ <u>1</u> t <u>.19</u> DC <u>.41</u>	Estimate <u>.027</u> Error <u>.18</u> WMNSQ <u>.99</u> t <u>0</u> DC <u>.42</u>
identifizieren			
Estimate <u>-1.14</u>	06 37	Estimate <u>-1.58</u>	20 Estimate <u>-.69</u>
Error <u>.19</u>		Error <u>.2</u>	45 Error <u>.17</u>
WMNSQ <u>.98</u>		39 WMNSQ <u>1</u>	53 WMNSQ <u>.99</u>
t <u>-.18</u>		48 t <u>-.75</u>	67 t <u>-.04</u>
DC <u>.44</u>		DC <u>.47</u>	69 DC <u>.44</u>
			72
integrieren			
Estimate <u>.12</u>	29 50 66 70	Estimate <u>.51</u>	01 Estimate <u>-.36</u>
Error <u>.18</u>		04 Error <u>.18</u>	10 Error <u>.20</u>
WMNSQ <u>1</u>		22 WMNSQ <u>1</u>	17 WMNSQ <u>.99</u>
t <u>.15</u>		36 t <u>.12</u>	63 t <u>0</u>
DC <u>.40</u>		46 DC <u>.4</u>	DC <u>.40</u>
		52	
generieren			
Estimate <u>.31</u>	28 38 40	Estimate <u>.03</u>	Estimate <u>.93</u>
Error <u>.16</u>		09 Error <u>.16</u>	Error <u>.17</u>
WMNSQ <u>1.01</u>		11 WMNSQ <u>1.02</u>	54 WMNSQ <u>1</u>
t <u>.34</u>		41 t <u>.55</u>	62 t <u>.05</u>
DC <u>.43</u>		42 DC <u>.41</u>	DC <u>.44</u>
		64	

8. Eidesstattliche Erklärung

Ich versichere, dass ich die Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen oder anderen Quellen entnommen sind, wurden als solche eindeutig kenntlich gemacht.

Die Arbeit ist noch nicht veröffentlicht und noch nicht als Dissertation zur Anerkennung oder Bewertung vorgelegt worden, weder von mir noch von anderen Doktoranden oder Studierenden an der TU Darmstadt oder an anderen Universitäten.

Ich bin mir bewusst, dass eine unwahre Erklärung rechtliche Folgen hat.

Darmstadt, 28.10.2014

Aileen Balkenhol

Aileen Deborah Balkenhol